



HAL
open science

Assimilation de données et méthodes adjointes pour la géophysique

Arthur Vidard

► **To cite this version:**

Arthur Vidard. Assimilation de données et méthodes adjointes pour la géophysique. Optimisation et contrôle [math.OA]. Université de Grenoble, 2012. tel-00939130

HAL Id: tel-00939130

<https://theses.hal.science/tel-00939130>

Submitted on 30 Jan 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Habilitation à Diriger les Recherches:

**Assimilation de données et méthodes adjointes pour
la géophysique**

Arthur Vidard

Table des matières

1	DES MÉTHODES DÉLICATES	9
1.1	L'assimilation de données dans un cadre d'océanographie opérationnelle	14
1.1.1	De l'importance de l'océanographie opérationnelle	14
1.1.2	Filtrage vs Contrôle	15
1.1.3	NEMOVAR et le projet VODA	19
1.2	Operateurs de covariances d'erreurs	19
1.2.1	Le rôle de \mathbf{B}	20
1.2.2	Corrélations monovariées	21
1.2.3	Corrélations multivariées	25
1.3	Le contrôle d'erreur modèle en assimilation variationnelle	27
1.4	Quelques pistes pour l'avenir	29
1.4.1	Evolution des méthodes et adaptation à l'évolution des moyens de calcul	29
1.4.2	Assimilation pour les systèmes de modélisation	29
1.4.3	Évolution de la formulation des statistiques d'erreurs	30
1.4.4	Méthodes hybrides et nouvelles méthodes	30
2	DES OBSERVATIONS RÉCALCITRANTES	33
2.1	Données altimétriques	36
2.1.1	Topographie Dynamique Moyenne	36
2.1.2	Tendance haussière du niveau de la mer	39
2.1.3	Une nouvelle génération	40
2.2	Assimilation d'images	41
2.2.1	Motivations	41
2.2.2	Images: Level(s) of Interpretation	43
2.2.3	Current Use of Images in DA: Pseudo Observation	45
2.2.4	Direct Assimilation of Images	50
2.3	Conclusions et perspectives de ce chapitre	57
3	DES MODÈLES INCOMPRIS	59
3.1	Étude de l'incrément d'analyse	59
3.2	Analyse de sensibilité	60
3.2.1	Analyse globale	61
3.2.2	Analyse locale	62
3.2.3	De l'analyse locale à l'analyse globale	64
3.2.4	Propagation d'incertitude	65

3.3	Estimation de paramètres et Contrôle robuste	66
3.4	Analyse de stabilité	68
3.5	Conclusion	70

INTRODUCTION

Habilitation à diriger les recherches... quel est le sens de cette démarche ? Un parcours rapide d'une sélection de tels documents ne m'a pas permis de dégager une structure et un type de contenu commun. Ma démarche ici sera de tenter de répondre à la question que soulève l'intitulé de ce diplôme : ai-je les qualités nécessaires pour diriger les recherches ? Les qualités humaines essentielles dans ce cadre sont difficiles à mettre en évidence dans un document écrit. Pour cela il serait plus pertinent de demander leurs avis aux personnes dont j'ai eu le plaisir d'encadrer les travaux (post-docs, ingénieurs et doctorants) ; on peut d'ailleurs regretter que le processus d'habilitation ne permette pas ce genre d'approche. Reste alors le contenu scientifique : une qualité qui me semble essentielle est le recul. Recul sur l'état de l'art dans le domaine considéré, possiblement sur les domaines connexes voire plus lointains. Il est important de bien comprendre les enjeux, les difficultés (résolues ou non) et les grands axes de recherches futurs. C'est dans cette optique qu'a été rédigé ce document. Sur les trois thèmes abordés, j'ai tenté de présenter une vision assez globale, en insistant sur certains points que j'estime importants et en donnant des directions de recherches futures. Il y a, à dessein, peu de résultats dans ce document, mais nombre des références sont données pour le lecteur curieux.

Mes préoccupations de recherche se situent dans l'amélioration des systèmes de prévision en géophysique, principalement pour l'océan et le couplage océan-atmosphère. En effet la compréhension et la prévision de l'atmosphère et de l'océan sont d'un intérêt sociétal majeur pour notre vie quotidienne. Cela concerne en particulier la prévision météorologique générale au niveau d'une région, d'un pays ou d'un continent dont dépendent de grands pans de notre activité économique (pêche, agriculture, transports aériens ou maritimes, etc.). Des prévisions à une échelle spatiale et temporelle plus fine sont également nécessaires pour réguler la production énergétique éolienne ou encore à des fins de prévention d'événements violents ou potentiellement dangereux à proximité de lieux sensibles tels que les aéroports. Des simulations à des échelles de temps beaucoup plus grandes sont requises dans les modèles climatiques afin d'analyser l'évolution sur le long terme de paramètres tels que le vent, la température atmosphérique ou océanique, la direction ou l'amplitude des courants océaniques, la salinité, l'humidité, les nuages, les précipitations, l'eau du sol, ou la concentration de certaines espèces chimiques. Un préalable indispensable à la prévision est de pouvoir reconstituer au mieux, l'état de l'écoulement à un instant « initial ». Cependant, l'information dont on dispose sur les écoulements géophysiques (atmosphère, océan, eaux continentales) est hétérogène en nature, qualité et densité.

La première source d'information prend la forme des équations qui régissent les écoulements : les modèles. Les fluides géophysiques sont représentés par des « variables d'état ». En météorologie ce sont les trois composantes du vent, la température, l'humidité, la pression. En océanographie on a les mêmes variables avec la salinité remplaçant l'humidité. Ces équations sont déduites des lois de conservation de masse, d'énergie et des lois de la thermodynamique. Elles ont la désagréable propriété d'être non linéaires ce qui signifie des interactions entre les différentes échelles du mouvement, et ceci a lieu tant en temps qu'en espace. Une autre difficulté est ce qu'on appelle les « effets sous-maille », par exemple les cumulo-nimbus,

ces nuages d'orage en forme d'enclume, contiennent une quantité d'énergie fantastique, l'équivalent de plusieurs bombes thermonucléaires, or leurs dimensions caractéristiques, de l'ordre d'une dizaine de kilomètres, dans toutes les directions, sont inférieures à la taille du pavé élémentaire. Pour avoir des modèles énergétiquement cohérents il est nécessaire de paramétrer ces flux d'énergie ce qui introduit des sources d'erreurs dans les modèles, ceci en plus des erreurs liées à la discrétisation.

En océanographie les mêmes types de difficultés sont retrouvés avec en plus l'existence de côtes qui imposent des contraintes aux limites supplémentaires. En toute rigueur on ne devrait pas séparer la modélisation de l'océan et de l'atmosphère tant sont grandes les interactions énergétiques entre ces deux composantes de l'enveloppe fluide terrestre : les ouragans qui ravagent les Caraïbes et les côtes du Golfe du Mexique prennent naissance dans des anomalies de la température de l'océan dans le Golfe de Guinée. Cependant les dimensions des systèmes numériques globaux à considérer sont encore hors de portée des ordinateurs actuellement existants. Les modèles actuels, tant en météorologie qu'en océanographie ne sont pas figés, ils évoluent continument avec une meilleure connaissance des processus physiques et grâce au progrès des méthodes numériques.

Les modèles mathématiques sont importants pour la compréhension de la dynamique de l'atmosphère et de l'océan. Mais si ils étaient notre seule source d'information aucune prévision ne serait possible faute, notamment au manque de la connaissance d'une condition initiale cohérente. On dispose également d'observations de ces systèmes en nombre de plus en plus important, notamment grâce aux nombreux satellites d'observation qui croisent maintenant au large de notre planète. Ces observations sont souvent indirectes et incomplètes, et de ce fait ne fournissent pas non plus, à elles seules, une connaissance approfondie de l'état du milieu considéré.

Et pour finir, depuis les débuts effectifs de la prévision numérique (Von Neumann et Charney dans les années 50), on dispose de statistiques sur les champs des variables atmosphériques, leur variabilité, leur cohérence en temps et en espace. Cette information fait aussi partie des données qu'il faudra utiliser. Comme nous le verrons dans la suite, l'utilisation de ces statistiques est très importante pour améliorer la prévision.

Je présenterai donc, dans ce document, des méthodes permettant de combiner tout ou partie de ces informations afin d'améliorer la prévision et la connaissance du fonctionnement de ces systèmes. Ces méthodes se basent le plus souvent sur une théorie mathématique solide, mais les appliquer dans un contexte réaliste n'est pas toujours chose aisée. C'est pourquoi je garde le souci d'accompagner les développements que nous effectuons jusqu'à des applications opérationnelles ou quasi opérationnelles afin de démontrer la faisabilité de ceux-ci. L'utilisation des méthodes de contrôle ou d'assimilation de données dans des contextes réalistes en géophysique demande de très lourds développements mais c'est le prix à payer pour accélérer le transfert des milieux académiques vers l'opérationnel. Il a fallu compter près de 20 ans pour que les méthodes adjointes proposées par Le Dimet dans le début des années 80 soit utilisés dans les systèmes de prévisions météorologiques opérationnelles et plus de 10 ans pour que soit ajoutée la prise en compte de l'erreur modèle pourtant déjà présente dans plusieurs travaux (dont ma thèse) au début des années 2000. Il me paraît important que dans une équipe de mathématiques appliquées on soit capable de mettre en œuvre une partie des méthodes développées dans des cas réalistes en lien avec des spécialistes du domaine considéré pour accélérer cette transmission et ne pas se contenter d'applications très idéalisées.

Ce document est divisé en trois chapitres, sur les trois acteurs principaux d'un système de prévision. Le premier s'intéresse aux méthodes d'assimilation de données proprement dites. Elles sont présentées principalement sous l'angle de l'approche par contrôle optimal, mais des liens et comparaisons avec les méthodes de filtrages sont également fournis. Un état de l'art est discuté avec un accent particulier sur les problèmes actuels, puis quelques directions futures sont évoquées. Le deuxième chapitre évoque le cas des observations au travers de deux cas concrets, un dont le traitement paraît simple mais qui ne l'est pas, un autre qui semble plus compliqué et qui l'est vraiment. Là également des perspectives sont présentes, sur les

améliorations nécessaires des méthodes et le traitement de nouvelles données. Le dernier chapitre est plus prospectif en ce qui me concerne, et décrit ce que les outils variationnels peuvent apporter au développement des modèles.

Chapitre 1

DES MÉTHODES DÉLICATES

Pour réaliser une prévision, le principe est simple : à partir d'une condition initiale du modèle on effectue une intégration numérique de celui-ci dont le résultat sera la prévision. Cependant nous ne disposons pas d'une connaissance suffisamment précise de cette condition initiale, mais seulement d'une estimation de celle-ci provenant, la plupart du temps, d'une prévision précédente. Les méthodes d'assimilation de données sont destinées à reconstituer cet état initial à partir des informations disponibles, à savoir les observations, les équations physiques (sous forme de modèles) et les statistiques d'erreur sur ces deux sources. Le problème est ardu : la condition initiale a une dimension de plusieurs centaines de millions alors que les observations ne se comptent qu'en millions. Mathématiquement on se retrouve donc dans le cadre d'un problème mal posé au sens de Tykhonov.

Tout l'art de l'assimilation de données réside donc dans la combinaison de toutes ces sources d'information hétérogènes. On peut classer les différentes méthodes actuellement considérées en deux branches principales : - Les méthodes basées sur le filtrage, et principalement le filtrage de Kalman étendu à des problèmes non linéaires. Une difficulté provient du fait que le filtrage nécessite la manipulation de matrices de covariances, avec les modèles opérationnels la taille des variables d'état ne permet pas l'utilisation de cette méthode directement et des versions sous-optimales ont dû être développées. De nombreuses variantes de ces filtres existent maintenant et sont devenues très performantes grâce à l'avènement des méthodes d'ensemble. - Les méthodes fondées sur le contrôle optimal (souvent appelées méthodes variationnelles). A chaque condition initiale possible du modèle on peut associer une trajectoire du modèle et calculer un écart à l'observation à l'aide d'une fonction quadratique. La meilleure condition initiale est celle qui minimise cette fonction. On obtient un problème bien posé en ajoutant un terme de régularisation qui contient l'information statistique sur les erreurs de prévision. Un avantage de cette méthode est de mettre en évidence un système d'optimalité dans lequel toute l'information est rassemblée. Du point de vue algorithmique la condition initiale est obtenue par la résolution d'un problème d'optimisation pour lequel de nombreux algorithmes robustes existent. Cette méthode, introduite en France au début des années 80, est maintenant utilisée par les principaux centres opérationnels dans le monde.

Il ne faut pas perdre de vue que la prévision gagne en valeur si on est capable de lui associer un indice de qualité ou une barre d'erreur. Ce dernier point est actuellement un axe important de recherche notamment par l'utilisation de techniques dites « d'ensemble » fondées sur des prévisions multiples à partir d'une série de conditions initiales ou aux frontières perturbées (ou parfois même avec des termes stochastiques dans les équations du modèle).

Par contrôle optimal

Les principes de base de l'Assimilation Variationnelle de Données (VDA) sont les suivants : L'état du système est représenté par la *variable d'état* notée $\mathbf{x}(t)$ fonction du temps et de l'espace. Elle appartient à l'espace des variables d'état \mathcal{X} et est composée des variables pronostiques du modèle (*e.g.* vitesse des courants, température, élévation de la surface de la mer, salinité, concentration des espèces biologiques ou chimiques, ...). L'évolution de \mathbf{x} est décrite par le système différentiel :

$$\begin{cases} \frac{d\mathbf{x}}{dt} = \mathcal{M}(\mathbf{x}), & t \in [0, T], \\ \mathbf{x}(0) = \mathbf{x}_0, \end{cases} \quad (1.1)$$

La condition initiale $\mathbf{x}_0 \in \mathcal{X}$ est inconnue et dépend de l'espace. On suppose que si \mathbf{x}_0 est connu, le problème (1.1) a une unique solution $\mathbf{x}(\mathbf{x}_0, t)$, $t \in [0, T]$. $\mathbf{y}(t) \in \mathcal{O}$ sont des observations du système disponibles entre 0 et T . L'espace \mathcal{O} est appelé espace des observations. L'écart entre les observations et la variable d'état est défini au moyen de la *fonction coût* suivante :

$$J(\mathbf{x}_0) = \frac{1}{2} \int_0^T \|\mathbf{y}(t) - \mathcal{H}[\mathbf{x}(\mathbf{x}_0, t)]\|_{\mathcal{O}}^2 dt + \frac{1}{2} \|\mathbf{x}_0 - \mathbf{x}^b\|_{\mathcal{X}}^2, \quad (1.2)$$

où \mathcal{H} est appelé *Opérateur d'Observation*. Ce dernier est une application depuis l'espace des variables d'état vers l'espace des observations \mathcal{O} où la comparaison sera effectuée (voir chapitre 2 pour plus de détails). \mathbf{x}^b est l'ébauche (le premier estimé, ou *first guess*) de la minimisation. Le second terme est un terme de régularisation au sens de Tikhonov (Tikhonov, 1963). Il permet d'introduire des informations *a priori*. Il est important de noter que J mobilise deux normes différentes qui définissent la structure topologique des espaces fonctionnels \mathcal{X} et \mathcal{O} ; en pratique ces normes dérivent d'un produit scalaire et peuvent tenir compte des statistiques d'erreur en introduisant des matrices de covariances d'erreur comme proposé par Mahalanobis (1936). Par exemple, une version discrétisée de la fonction coût peut faire intervenir

$$\|\mathbf{x}_0 - \mathbf{x}^b\|_{\mathcal{X}}^2 := (\mathbf{x}_0 - \mathbf{x}^b)^* \mathbf{B}^{-1} (\mathbf{x}_0 - \mathbf{x}^b), \quad (1.3)$$

où \mathbf{B} est la matrice de covariance d'erreur d'ébauche. l'exposant * symbolise l'opérateur adjoint. On définit de la même façon la norme $\|\cdot\|_{\mathcal{O}}^2$ en utilisant la matrice \mathbf{R} de covariance d'erreur d'observation.

Le problème d'assimilation de données peut être posé comme la recherche de l'état analysé \mathbf{x}_0^a qui minimise J vérifiant

$$\mathbf{x}_0^a = \operatorname{argmin} J(\mathbf{x}_0). \quad (1.4)$$

Une condition nécessaire pour que \mathbf{x}_0^a soit une solution de (1.4) est que

$$\nabla J(\mathbf{x}_0^a) = 0 \quad (1.5)$$

qui devient une condition suffisante si J est convexe. En première approximation on en revient à une minimisation sans contrainte malgré le fait que, par exemple, certaines quantités doivent être positives (humidité, salinité, concentration, ...). D'un point de vue pratique, \mathbf{x}_0^a est estimé au moyen d'un algorithme de descente, c'est à dire en trouvant la limite de la suite

$$\mathbf{x}_0^{k+1} = \mathbf{x}_0^k + \lambda_k D_k, \quad (1.6)$$

où λ_k est le pas réalisant le minimum de J suivant la direction de descente D_k déduite du gradient de J . Afin de calculer le gradient ∇J par rapport à la variable de contrôle \mathbf{x}_0 , on introduit la *variable adjointe* \mathbf{p}

solution du modèle adjoint :

$$\begin{cases} \frac{d\mathbf{p}}{dt} + \left[\frac{\partial \mathcal{M}}{\partial \mathbf{x}} \right]^* \cdot \mathbf{p} = \left[\frac{\partial \mathcal{H}}{\partial \mathbf{x}} \right]^* \cdot (\mathcal{H}(\mathbf{x}) - \mathbf{y}), & t \in [T, 0], \\ \mathbf{p}(T) = 0. \end{cases} \quad (1.7)$$

Après une intégration rétrograde du modèle adjoint, le gradient de J est donné par :

$$\nabla J = -\mathbf{p}(0) \quad (1.8)$$

La dérivation du système (1.7) peut-être trouvée dans [Le Dimet and Talagrand \(1986\)](#). Le modèle (1.1) plus le modèle adjoint (1.7) accompagné de la fonction coût (1.2) forment ce qu'on appelle le *Système d'optimalité*. La variable adjointe a la même dimension que la variable d'état du modèle direct.

Dans le cas où \mathcal{M} et/ou \mathcal{H} sont non-linéaires, afin d'éviter d'avoir à minimiser une fonction non convexe, l'algorithme 4D-Var incrémental a été proposé ([Courtier et al., 1994](#)). On résout alors le même problème que le 4D-Var par approximation successives basées sur ladite *Approximation Linéaire Tangente* (TLA) :

$$\mathcal{M}_{0 \rightarrow i}(\mathbf{x}_0) - \mathcal{M}_{0 \rightarrow i}(\mathbf{x}_0^b) \simeq \mathbf{M}_{0 \rightarrow i} \cdot (\mathbf{x}_0 - \mathbf{x}_0^b) \quad (1.9)$$

et

$$\mathcal{H}_i(\mathbf{x}_i) - \mathcal{H}_i(\mathbf{x}_i^b) \simeq \mathbf{H}_i \cdot (\mathbf{x}_i - \mathbf{x}_i^b)$$

Avec $\mathbf{M} = \frac{\partial \mathcal{M}}{\partial \mathbf{x}}$ et $\mathbf{H} = \frac{\partial \mathcal{H}}{\partial \mathbf{x}}$.

En définissant l'incrément $\delta \mathbf{x}_0$ comme

$$\delta \mathbf{x}_0 = \mathbf{x}_0 - \mathbf{x}_0^b \quad (1.10)$$

on peut réécrire la fonction coût du 4D-Var comme une fonction de $\delta \mathbf{x}$ équivalente :

$$J(\delta \mathbf{x}_0) = \frac{1}{2} \int_0^T \|\mathbf{y}(t) - \mathcal{H}[\mathbf{x}(\mathbf{x}^b + \delta \mathbf{x}_0, t)]\|_{\mathcal{O}}^2 dt + \frac{1}{2} \|\delta \mathbf{x}_0\|_{\mathcal{X}}^2, \quad (1.11)$$

Si on définit le *vecteur innovation* \mathbf{d}_i tel que

$$\mathbf{d}_i = \mathbf{y}_i^o - \mathcal{H}_i(\mathcal{M}_{0 \rightarrow i}(\mathbf{x}^b)) \quad (1.12)$$

et en utilisant la TLA, on peut proposer l'approximation de J :

$$\tilde{J}(\delta \mathbf{x}_0) = \frac{1}{2} \int_0^T \|\mathbf{d}(t) - \mathbf{H}\mathbf{M}_{0 \rightarrow i}\delta \mathbf{x}_0\|_{\mathcal{O}}^2 dt + \frac{1}{2} \|\delta \mathbf{x}_0\|_{\mathcal{X}}^2, \quad (1.13)$$

$\tilde{J}(\delta \mathbf{x})$ est maintenant une fonction quadratique et a donc un minimum unique.

Au cours de la minimisation, $\delta \mathbf{x}$ va devenir trop important, rendant caduque la TLA. Afin de pallier cette incohérence, on arrête la minimisation, actualise les opérateurs linéarisés \mathbf{M}_i et \mathbf{H}_i et le vecteur innovation \mathbf{d} en recalculant la trajectoire non linéaire à partir de la condition initiale corrigée $\mathbf{x}_0 = \mathbf{x}_0^b + \delta \mathbf{x}_0$, et on itère ce processus plusieurs fois. Les minimisations quadratiques sont habituellement appelées boucles internes, tandis que les étapes de relinéarisation sont appelées itérations externes.

Cette façon de résoudre un problème des moindres carrés non linéaire est appelée l'algorithme de Gauss-Newton dans la communauté des spécialistes d'optimisation. Il converge si l'ébauche est assez proche du minimum global ('assez' dépendant de la forme de la fonction coût elle même) ([Nocedal and Wright, 1999](#)).

Il est important de noter que, contrairement à ce qui est souvent dit, le 4D-Var standard ne requiert pas la validité de cette TLA. Le calcul du gradient par méthode adjointe reste exact même en cas d'opérateurs non linéaires. Cependant la présence de non linéarité induit celle de minimas locaux et peut empêcher la convergence globale de la minimisation.

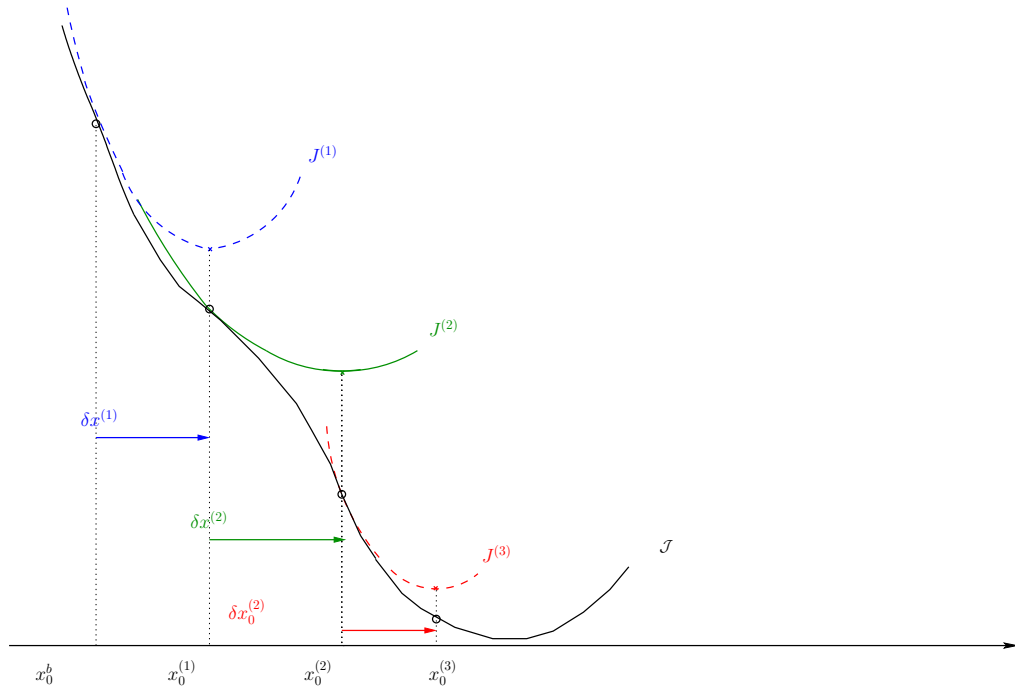


FIGURE 1.1: Représentation schématique de l'algorithme incrémental (Gauss-Newton). En noir la fonction coût non quadratique, en couleur les approximations quadratiques successives.

Multi-Incremental: On peut utiliser une autre approximation que la TLA dans ce contexte. Les centres de prévisions météorologiques (Météo-France, ECMWF, MetOffice (UK), ...) utilisent l'approche Multi-incrementale (Veersé and Thépaut, 1998) où le modèle utilisé dans les minimisations successives de la boucle interne sont des approximations du modèle tangent (résolution plus basse et physique simplifiée, approximation non linéaire pour le MetOffice). La première itération externe est réalisée en utilisant une résolution grossière dans la boucle interne, et la résolution (et la physique) est améliorée pour chacune des itérations externes suivantes.

Les hypothèses successives sont :

$$\mathcal{M}_{0 \rightarrow i}(\mathbf{x}_0^b + \mathbf{S}^{-I} \delta \mathbf{x}_0^s) - \mathcal{M}_{0 \rightarrow i}(\mathbf{x}_0^b) \simeq \mathbf{S}^{-I} \mathbf{M}_{0 \rightarrow i}^s \delta \mathbf{x}_0^s \quad (1.14)$$

l'exposant s signifiant état/modèle simplifié et \mathbf{S}^{-I} est l'inverse généralisée (ou pseudo inverse) de \mathbf{S} .

L'opérateur de simplification \mathbf{S} est différent d'une itération externe à l'autre (de la plus simple à la plus réaliste). Cette approche a l'avantage de fortement réduire le coût de la minimisation, une boucle interne à résolution complète étant beaucoup trop coûteuse à réaliser dans les systèmes de prévision numérique du temps actuels, de plus cela permet de simplifier les parties du modèle qui ne sont pas différentiables. Cependant, contrairement à l'algorithme Gauss-Newton (ou incrémental), il n'y a pas de résultat de convergence et en pratique de tels algorithmes finissent par diverger après un petit nombre d'itérations externes (Trémolet, 2007b). Il faut cependant noter qu'en pratique ce n'est pas une limitation sévère puisque le nombre d'itération est de toute façon limité pour des raisons de coût de calcul. Dans sa thèse, Neveu (2011) a montré sur un exemple académique que l'algorithme multi-incremental (sans simplification de la physique) pouvait converger rapidement vers le minimum global durant les quelques premières itérations externes et diverger aussi rapidement durant les suivantes, rendant délicat le choix de ce nombre d'itérations externes. On peut voir le multi-incremental comme une méthode multi-échelle du pauvre. Neveu (2011) a

ainsi proposé d'utiliser deux méthodes multi-échelles proprement définies avec des propriétés de convergence connues (Full Approximation Scheme et Multi-Grid Gauss Newton). Bien que très prometteuses, ces approches devront probablement faire l'objet d'une investigation supplémentaire avant de pouvoir les appliquer à des configurations réalistes étant donné qu'elles nécessitent une minimisation à haute résolution (et même de la fonction coût non quadratique pour l'approche FAS). De plus l'algorithme multi-incremental converge de façon diabolique durant les premières itérations, il sera donc difficile de proposer des méthodes plus performantes.

Par filtrage

La deuxième grande classe de méthode d'assimilation de données est basée sur le filtrage. N'étant pas expert de ce type de méthode, je me contenterai d'en donner les grandes lignes. Les méthodes de filtrage en assimilation de données sont principalement basées sur le filtrage de Kalman. On considère alors l'évolution temporelle du modèle comme un processus aléatoire discret, et l'on corrige les prévisions du modèle en tenant compte des observations au fur et à mesure de leur disponibilité. Le filtre de Kalman fournit un algorithme pour le calcul d'estimés statistiquement optimaux pour les modèles linéaires (Kalman, 1960). Pour les modèles faiblement non linéaires son algorithme peut être modifié par un développement en série de Taylor local autour de l'estimé courant. Le filtre de Kalman étendu est un exemple d'approximation du filtre au premier ordre. De nombreux algorithmes basés sur ce principe ont été présentés dans la littérature (par exemple Jazwinski (1970)). Comme le coût en calcul augmente rapidement avec l'ordre du développement, le nombre de publications décroît aussi rapidement avec la précision de l'approximation.

Sous quelques hypothèses de normalité et de décorrélation, on obtient les équations suivantes pour le filtre de Kalman "étendu" (EKF), *i.e.*:

$$\begin{aligned}
\mathbf{x}_{n+1}^f &= \mathcal{M}_{n \rightarrow n+1}(\mathbf{x}_n^f) \\
\mathbf{P}_{n+1}^f &= \mathbf{M}_{n \rightarrow n+1}(\mathbf{x}_n^a) \mathbf{P}_n^a \mathbf{M}_{n \rightarrow n+1}^T(\mathbf{x}_n^a) + \mathbf{Q}_n \\
\mathbf{K}_{n+1} &= \mathbf{P}_{n+1}^f \mathbf{H}_{n+1}^T \mathbf{x}_{n+1}^f \left[\mathbf{H}_{n+1}(\mathbf{x}_{n+1}^f) \mathbf{P}_{n+1}^f \mathbf{H}_{n+1}^T \mathbf{x}_{n+1}^f + \mathbf{R}_{n+1} \right]^{-1} \\
\mathbf{P}_{n+1}^a &= \mathbf{P}_{n+1}^f - \mathbf{K}_{n+1} \mathbf{H}_{n+1}(\mathbf{x}_{n+1}^f) \mathbf{P}_{n+1}^f \\
\mathbf{x}_{n+1}^a &= \mathbf{x}_{n+1}^f + \mathbf{K}_{n+1} (\mathbf{y}_{n+1}^o - \mathcal{H}_{n+1}(\mathbf{x}_{n+1}^f))
\end{aligned} \tag{1.15}$$

\mathbf{P}^f , \mathbf{P}^a et \mathbf{Q} étant les matrices de covariance d'erreur de prévision, d'analyse et de modèle respectivement. Les trois étapes du milieu sont couteuses.

Le gros désavantage du filtre de Kalman ainsi défini est que pour des problèmes de hautes dimensions, le fardeau en coût de calcul s'alourdit rapidement pour devenir impraticable. Récemment une très forte activité de recherche autour des méthodes issues du filtre de Kalman a pu être recensée et a permis une grande avancée et un foisonnement des variantes de celui ci. Parmi celles ci on peut mentionner :

- Le filtre de Kalman d'Ensemble (Evensen (1994), Evensen (2003)). Cet algorithme basé sur une approche de type Monte-Carlo calcule les covariances d'erreur de prévision (deuxième étape de 1.15) en intégrant en temps un ensemble dont les états initiaux et les forçages ont été perturbés aléatoirement. L'analyse des états perturbés est effectuée dans les filtres les plus utilisés, en utilisant des observations également perturbées.
- Les filtres racine carré de rang réduit (Pham et al., 1998; Verlaan and Heemink, 1997; Brasseur, 2006; Brasseur and Verron, 2006; Brankart et al., 2010, 2011) sont des algorithmes basés sur une décomposition en valeurs propres des matrices de covariance d'erreur. Les vecteurs propres associés aux plus grandes valeurs propres de la matrice de covariance d'erreur d'analyse sont utilisés pour former un ensemble de vecteurs d'état perturbés qui sont ensuite intégrés en temps par le modèle non

linéaire. Cet ensemble a la propriété que dans le cas d'un modèle linéaire, les covariances dans le sous espace engendré par les prévisions de l'ensemble sont exactes. Dans le cas d'un modèle non linéaire, la méthode est approchée et comparable au filtre de Kalman étendu.

Les relations entre le EnKF et les autres filtres de rang réduit sont bien discutées dans [Nerger et al. \(2005\)](#) qui appellent ESKF (Error Subspace Kalman Filters) ces formes d'alternative à l'EKF. On reprendra cette appellation dans le reste du document. D'autres alternatives sont disponibles dans la littérature, comme par exemple les filtres du second ordre qui gardent un terme de plus dans le développement de Taylor et ensuite appliquent une des solutions ci dessus ([Pham, 1996, 2001](#); [Verlaan and Heemink, 2001](#))

1.1 L'assimilation de données dans un cadre d'océanographie opérationnelle

L'importance clé de l'assimilation de données pour l'océanographie et en particulier pour l'océanographie opérationnelle est maintenant bien reconnue. Un large spectre de recherches et d'applications dépend de la disponibilité rapide de produits informatifs sur l'océan. Le domaine de l'océanographie physique a maintenant atteint une maturité suffisante pour permettre de combiner les informations issues des modèles numériques et celles issues des observations au moyen des méthodes d'assimilation de données. Cela permet de fournir des analyses et des prévisions océaniques pour diverses échelles temporelles et spatiales.

1.1.1 De l'importance de l'océanographie opérationnelle

Le développement de l'océanographie opérationnelle est relativement récent comparé à celui de sa grande soeur atmosphérique et surtout ne dispose pas des mêmes moyens. Cependant comprendre et prédire la dynamique océanique est actuellement le sujet d'un intense effort de recherche par la communauté scientifique internationale. Cet effort est motivé par deux principaux moteurs : 1. Le rôle crucial de l'océan dans la détermination du climat, en particulier dans la perspective de son changement global. L'océan joue aussi un rôle essentiel pour les prévisions météorologiques aux échelles de temps saisonnière à décennales, avec des applications importantes dans la prévision des sécheresses, l'énergie, l'agriculture, etc. 2. Le développement de l'océanographie opérationnelle, c'est à dire la prévision en temps réel de la circulation océanique, avec des applications dans le trafic maritime, la pêche, la prévision du temps, la pollution, la gestion des écosystèmes, etc.

En France cet essor est caractérisé par la création de Mercator-Océan qui prend racine dans la deuxième moitié des années 1990 (1994, <http://www.mercator-ocean.fr/>, [Brasseur et al. \(2005\)](#)). Au niveau Européen cet effort a été fédéré au travers du Global Monitoring for Environment and Security (GMES <http://www.gmes.info/>) et de sa composante d'océanographie opérationnelle MERSEA (2003-2008, <http://www.mersea.eu.org/>) puis MCS/MyOcean (2008- <http://www.gmes.info/175.0.html>), et au niveau international par le programme de l'UNESCO GODAE (1997-2007 <http://www.godae.org/>) qui vise à développer un système global de surveillance et de prévision aux échelles globales et régionales.

Dès 1998, les cartes de la topographie des océans tracées par Topex/Poséidon et ERS (projet Ssalto/Duacs) sont pour la première fois élaborées et diffusées en temps réel. L'océan profond, qui ne peut s'observer qu'*in situ*, prend le relais suivant avec le programme international Argo et sa composante française Coriolis. Le déploiement de 3000 flotteurs à travers l'océan global est entrepris pour apporter aux modèles les indispensables mesures verticales de température et salinité, entre 0 et 2000 m de profondeur (voir chapitre 2 pour plus de considérations sur les observations).

1.1.2 Filtrage vs Contrôle

Les modèles et les observations ayant atteint une maturité suffisante pour effectuer de la prévision, il est nécessaire de mettre en place des systèmes d'assimilation pour initialiser les modèles. Les techniques d'assimilation de données actuellement ou jusqu'à très récemment utilisées dans les centres opérationnels fournissant des prévisions océanographiques (par exemple Mercator Océan, ECMWF, ...) sont basées sur des implémentations sous optimales du filtre de Kalman (Interpolation optimale (OI), Filtres de Kalman Etendus de rand réduit (e.g., SEEK, Brasseur and Verron (2006)) et Filtres de Kalman d'Ensemble) ou sur des méthodes variationnelles de type 3D-Var / 3D-Fgat (Marine Forecast System de l'INGV, nouveau système à l'ECMWF et le UK-MetOffice, ...). En météorologie l'OI a longtemps été utilisée pour produire des conditions initiales pour la prévision numérique du temps (NWP; Lorenc (1981)). La plupart des principaux centres de NWP ont maintenant remplacé leur système basé sur l'OI par des systèmes de type 4D-Var (Rabier, 2005).

En raison de ressources plus limitées comparées à la prévision numérique du temps, l'océanographie opérationnelle n'a pas basculé vers le 4D-Var au même moment, ce type d'algorithme requérant un effort d'ingénierie et de calcul important. Le choix des futurs systèmes d'assimilation pour l'océan est toujours un débat ouvert (et parfois tendu), l'alternative offerte par les ESKF ayant connu un fort développement ces dernières années et proposant ainsi une alternative crédible pour les systèmes opérationnels.

Le choix amené à faire par les centres d'océanographie opérationnels entre ces deux types de méthode n'est pas évident, et n'est sans doute pas unique. On peut tout de même essayer de dresser un portrait de leurs forces et faiblesses respectives. Ce portrait sera forcément un peu biaisé étant donné que je suis plutôt un spécialiste des méthodes basées sur le contrôle.

Les ESKFs ont plusieurs avantages majeurs, le premier d'entre eux, qui a sans doute contribué à la forte intensité de recherche à leur sujet, et que ces filtres sont relativement faciles à implémenter, notamment parce qu'ils ne nécessitent pas la dérivation des modèles tangent et adjoint. Nerger et al. (2005) déclare même que cette implémentation ne nécessite que quelques centaines de lignes pour les plus compliquées (filtres du second ordre, en utilisant des bibliothèques mathématiques classiques). Même si ils font là fi des réglages qui peuvent être délicats et si un système opérationnel demandera sans doute plus de complexité pour gérer certaines spécificités, on comprend tout de même qu'il est aisé de mettre en place une telle approche, permettant ainsi d'utiliser les ressources forcément limitées à d'autres objectifs. Un autre atout très important pour les années à venir est la parallélisation "naturelle" de ces algorithmes : l'évolution par le modèle des différents membres de l'ensemble est indépendante. Alors que les moyens de calculs s'orientent de plus en plus vers des solutions massivement parallèles, c'est un argument majeur qui peut peser dans la balance du choix d'un système opérationnel. Les deux premières qualités qui viennent d'être citées sont essentiellement d'ordre technique et ne sauraient suffire à justifier l'emploi de telles méthodes. Les ESKFs ont également des avantages scientifiques certains. Premièrement, ce type d'algorithme fournit de façon naturelle une estimation des statistiques d'erreur d'analyse et de prévision qui évoluent au fur et à mesure de l'assimilation des observations et par conséquent sont dépendantes de l'état courant du système (lignes 1 et 3 de l'équation 1.15). Un autre élément clé en faveur des méthodes d'ensemble dérivées du filtre de Kalman est qu'elles fournissent naturellement un ensemble pertinent de conditions initiales pour les systèmes de prévision d'ensemble (EPS) couramment utilisées dans les principaux centres de prévision météorologiques pour modéliser les incertitudes sur la prévision (voir par exemple Molteni et al. (1996)).

Comme on vient de le voir les dérivées du filtre de Kalman étendu présentent de nombreux avantages séduisants, tout n'est cependant pas rose. Selon le type choisi, ces algorithmes demandent à certaines étapes une approximation linéaire du modèle et/ou de l'opérateur d'observation ou une approximation gaussienne des statistiques d'erreur obtenues. En allant vers une plus haute résolution et des observations de plus en

plus indirectes, il est vraisemblable que ces approximations deviennent trop fortes et dégradent la qualité des filtres. De plus, pour des raisons de coût, les méthodes séquentielles assimilent les observations “en paquet” (regroupant plusieurs temps différents en un seul), rendant leur comparaison avec le modèle approximative. Dans le cas des filtres racine carrée, il est cependant possible de contrevenir à cette limitation par l’utilisation de méthodes de type FGAT (voir plus loin) au prix d’une intégration supplémentaire du modèle. Le plus grand problème potentiel pour de futures applications opérationnelles de ce type de méthodes réside, à mon avis, dans la taille forcément réduite des ensembles.

Les ESKF sont significativement plus performants que l’EKF traditionnel pour des systèmes chaotiques tels que l’attracteur de Lorenz souvent utilisé comme cas test, signifiant ainsi que non seulement ils fournissent une alternative accessible pour les problèmes de grande taille, mais ils permettent une meilleure prise en compte des non linéarités. Cependant pour une application sur des systèmes multi échelles comme le sont les modèles atmosphériques et océaniques des approximations supplémentaires doivent être faites. Quand on assimile les données sur des systèmes de ce type, la dimension du vecteur d’état discrétisé n est plusieurs ordres de grandeur supérieur à la taille de l’ensemble N compte tenu de leur coût en calcul (N typiquement de l’ordre de 100 pour l’EnKF). Ceci engendre des corrélations irréalistes dans la matrice de covariance générée par l’ensemble et par conséquent une influence trop lointaine des observations dans les incréments créés par le filtre. Ce problème peut être grandement diminué par des ‘localisations de covariance’ qui suppriment artificiellement ces effets distants indésirables, en utilisant une fonction d’atténuation (un produit de Schur entre les matrices de covariance générées par l’ensemble et une fonction de covariance à support compact ou une analyse par boîte (Ott et al., 2004)) au risque d’en supprimer d’autres parfaitement légitimes. Cependant la localisation est une modification supplémentaire des équations du filtre de Kalman qui peut potentiellement détruire les équilibres physiques naturellement conservés par le filtre. La résolution spatiale des grilles de discrétisation augmentant, la taille trop réduite de l’ensemble risque de rendre les statistiques déduites moins significatives et de ce fait engendrer des besoins en localisation trop importants. On n’en est pas encore arrivé à ce stade et on pourrait imaginer une approche multi-résolution où, à l’instar du 4D-Var multi-incrémental, un ensemble à plus basse résolution est utilisé pour l’analyse haute résolution. Ce type d’approche a d’ailleurs déjà été proposé par Gao and Xue (2008) mais cela introduit une nouvelle source d’erreur qu’il conviendra alors d’étudier attentivement. Les aspects “rang réduit” induisent aussi potentiellement une dégénérescence du filtre. On a alors recours à des techniques dites d’inflation des covariances, qui si elle se justifie en pratique, sont plus questionnables théoriquement.

L’assimilation variationnelle permet de dépasser les nombreuses limitations des implémentations sous optimales du filtre de Kalman: elle permet une plus grande flexibilité dans l’assimilation de différents types d’observation quelque soient leur distribution spatiale et temporelle et leur caractère indirect. Il n’est pas nécessaire d’effectuer de localisation des observations, leur permettant ainsi, en principe, d’influencer l’analyse en tous points de grille du modèle. De façon plus importante, elle fournit un cadre plus général pour employer des représentations plus sophistiquées (de rang complet) des modèles de covariance d’erreur, telle, par exemple qui pourraient prendre en compte des erreurs corrélées en temps ou des erreurs d’observations non gaussiennes, bien que des densités de probabilité non gaussiennes pour les erreurs d’observation ou d’ébauche rendront la fonction coût non quadratique. Dans le cas de modèles et d’opérateurs d’observation non linéaires, les méthodes variationnelles ne font pas d’approximation linéaire. Même si la méthode de minimisation (4Dvar incrémental) peut employer de telles approximations, à convergence, si la méthode converge, le minimum obtenu est bien celui du problème non linéaire. De plus, pour aider à cette convergence, des méthodes multi échelles solides existent pour résoudre la minimisation même à haute résolution, celles-ci ne sont cependant peu ou pas encore employées (la méthode multi-incrémentale peut-être considérée comme une méthode multi-échelle dont il manquerait quelques étapes).

Cependant elle a aussi quelques désavantages clés. Sans doute le plus significatif est que l’algorithme,

au point de convergence, fournit un état estimé optimal, mais ne produit pas de manière claire d'estimés actualisés des statistiques de l'erreur d'analyse et un effort supplémentaire et non trivial doit être fait pour les obtenir (Gejadze et al., 2011). Cela peut paraître anecdotique, mais cette information est très précieuse afin de pouvoir faire évoluer la matrice d'erreur débauche de cycle temporel en cycle temporel. De plus, pour des systèmes non linéaires, la fonction coût à minimiser devient non convexe et fait ainsi apparaître des minimas locaux. Le nombre de ces minimas locaux augmente, à système donné, avec la résolution et la taille de la fenêtre d'assimilation. Ce dernier point est illustré sur la figure 1.2 qui montre, pour deux tailles de fenêtre d'assimilation, la forme de la fonction coût d'une variable de contrôle du système de Lorenz. Ceci peut rendre l'algorithme extrêmement sensible à l'ébauche de la minimisation. En effet les méthodes de gradient ne garantissent que la convergence vers des minimas locaux, ainsi si l'ébauche est localisée dans une région d'attraction d'un de ces minimas locaux, la solution du 4D-Var va converger vers un estimé sous optimal du contrôle. Ceci limite grandement la taille de la fenêtre temporelle d'assimilation. On peut également noter un autre aspect qui affecte la taille de la fenêtre d'assimilation : le conditionnement de la minimisation (et donc sa rapidité de convergence) est affecté par le nombre d'observations (Haben et al., 2011)

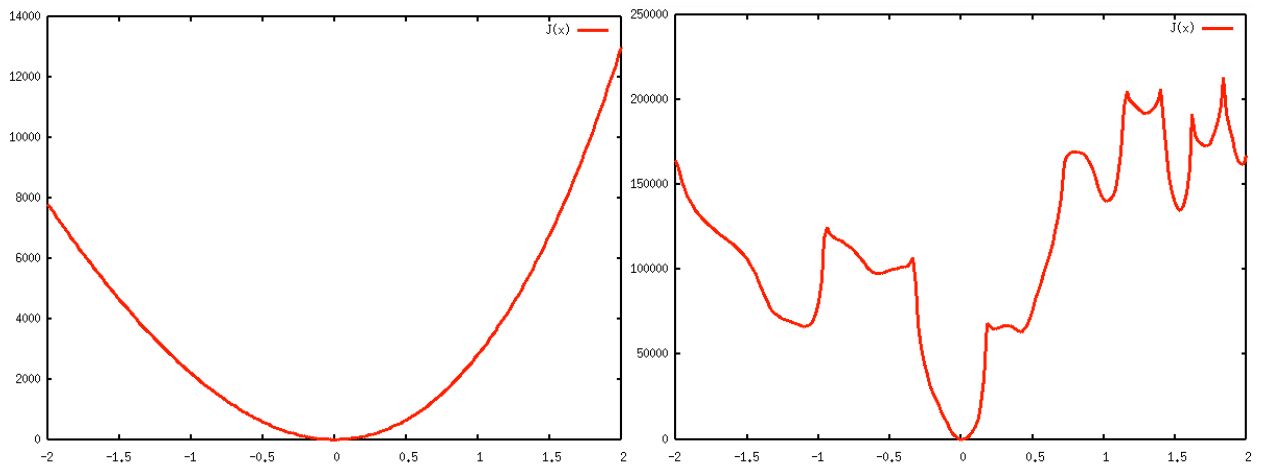


FIGURE 1.2: Valeur de la fonction coût autour de la valeur optimale pour un système de Lorenz avec des fenêtres d'assimilation de $T=20$ et $T=50$

Un autre aspect important en défaveur de l'assimilation variationnelle réside dans les développements très substantiels qu'elle requiert comparé aux autres méthodes, en particulier le développement des modèles linéaire tangent et adjoint du modèle dynamique océanique ainsi que des opérateurs de covariance d'erreurs efficaces.

Plusieurs études de comparaison entre les approches de type variationnelle et EnKF ont été publiées récemment (Fertig et al., 2007; Davey et al., 2006; Kalnay et al., 2007, ...), la plus convaincante (Buehner et al., 2010a,b) proposant une comparaison sur des systèmes opérationnels en réduisant au minimum les différences. La conclusion de cette étude est que, sur leur système (modèle atmosphérique 35 km de résolution), les deux approches fournissent une analyse de qualité équivalente pour un coût similaire (nous voilà bien avancés).

Cependant il n'est pas évident de tirer de ces études comparatives des conclusions définitives (mais il y en a-t-il vraiment ?). En effet, faire une comparaison rigoureuse est assez délicat, les approches et surtout les approximations et même les astuces couramment pratiquées sont différentes. Même dans un contexte favorable, il est difficile de faire exactement la même chose, ou alors, et c'est souvent ce qui est fait, c'est

sur des cas tests trop simples. On a souvent la tentation d'utiliser des modèles tels l'attracteur de Lorenz, ce qui nous donne une information sur la robustesse de la méthode à un comportement chaotique, mais quid de l'influence de la taille du problème ou de l'utilisation d'observations exotiques ? De plus ces comparaisons sont parfois faites par des spécialistes de l'une des approches comparées alors que l'implémentation et les réglages de ces méthodes sont délicats et demandent une expertise certaine.

Même dans le cas de [Buehner et al. \(2010a,b\)](#) où deux systèmes opérationnels sont comparés, ils l'ont été à caractéristiques fixées, je suppose pour obtenir un coût équivalent, mais n'aurait-on pas obtenu une conclusion différente avec plus d'itérations du 4Dvar ou plus de membres dans l'EnKF ?

Mon sentiment est qu'il faudrait un ensemble de cas test, comme il est en train de se mettre en place dans les projets [COMODO](#) (pour la modélisation et les schémas numériques) ou [SANGOMA](#) (pour comparer les approches de filtrages entre elles), pour évaluer la robustesse des méthodes aux différentes difficultés que les méthodes sont amenées à rencontrer (non linéarité, taille du problème, coût numérique, non différentiabilité...). Certaines des difficultés sont communes, en effet comme en 4DVar, les non linéarités limitent la taille de la fenêtre d'assimilation pour l'EKF. Si le 4Dvar s'améliore dans un premier temps avec l'allongement de la fenêtre, les ESKF se dégradent ([Fertig et al 2007](#)). Le contrôle d'erreur modèle atténue ce problème dans les deux cas.

Une forte activité de recherche a lieu actuellement pour circonvier aux problèmes précités des méthodes d'assimilation de données, on peut par exemple citer la possibilité de considérer des statistiques non gaussiennes dans les méthodes de filtrages par anamorphose ([Simon and Bertino, 2009](#)) et gaussiennes tronquées ([Lauvernet et al., 2009](#)) ou l'utilisation de méthodes primales-duales hautement parallélisables pour résoudre le 4D-Var (confère plus loin).

Comme je le sous-entends quelques lignes plus haut, il n'y a sans doute pas une réponse universelle sur quelle méthode choisir, cela dépend de l'application ciblée, des moyens humains et matériels dont on dispose et sans doute d'autres paramètres.

Récemment, de nouveaux acteurs ont fait leur apparition sur la scène de l'assimilation de données. Pour les systèmes non linéaires les filtres issus du filtre de Kalman sont sous optimaux notamment en raison de l'hypothèse de gaussianité de l'erreur d'analyse. La formulation plus générale des filtres particuliers (PF) propose de propager un ensemble de "particules" représentant plusieurs trajectoires potentielles du système, d'une manière très similaire à ce qui se fait pour les membres de l'EnKF. Dans les PF cependant chaque particule a un poids associé qui est utilisé pour calculer la moyenne et les moments d'ordre supérieurs. Contrairement à l'EnKF, aux temps des observations, celles ci sont utilisées pour actualiser le poids associé à chaque particule sans modifier celle-ci. Si l'on dispose d'un nombre infini de particules on peut montrer que cette approche est optimale même pour des systèmes non linéaires et des statistiques d'erreur non gaussiennes. Cependant le coût de ce genre de filtre est prohibitif pour les systèmes qui nous intéressent ici en raison du nombre important de particules nécessaire à leur bon fonctionnement. Adapter les filtres particuliers à l'assimilation de données en géophysique est actuellement un axe de recherche important, on peut par exemple citer le travaux de [Hoteit et al. \(2008a\)](#) [Bengtsson et al. \(2003\)](#); [Hoteit et al. \(2008b, 2011\)](#), sur le filtre de Kalman particulière (PKF) qui tente de combiner les PF et EnKF ou les travaux de [Van Leeuwen and Ades \(2012\)](#) sur une autre manière de calculer le poids de chaque particule. Outre les nouveaux filtres, on peut également citer le nudging direct et rétrograde (BFN, ([Auroux and BLUM, 2008](#))) où un terme de nudging (rappel newtonien vers les observations) est utilisé pour stabiliser des intégrations rétrogrades du modèle. Cet algorithme est très facile à mettre en place et semble donner une analyse de bonne qualité, qui peut ensuite être utilisée comme point de départ pour une méthode plus sophistiquée ([Auroux and BLUM, 2008](#)). Cet algorithme a été utilisé avec succès par [Boilley and Mahfouf \(2012\)](#) pour assimiler des données dans un modèle atmosphérique à haute résolution où la mise en pratique d'une méthode plus classique aurait été très délicate.

1.1.3 NEMOVAR et le projet VODA

Comme dit plus haut, un intense effort de recherche est en œuvre pour circonvenir aux inconvénients des deux grandes classes de méthodes par leur spécialistes respectifs (ou par les rare spécialistes œcuméniques); afin d’apporter ma pierre à l’édifice, j’ai mis en place le projet VODA (Variational Ocean Data Assimilation) qui a reçu le soutien de l’ANR et vise à bâtir un système d’assimilation de données variationnelle performant pour la recherche et l’opérationnel autour de la plateforme de modélisation de l’océan NEMO. Ce projet se base sur l’initiative NEMOVAR (assimilation variationnelle pour NEMO) dont je fais également partie, en collaboration avec l’ECMWF, le CERFACS et le MetOffice, qui a abouti à construire un système 3DFGAT (First Guess at Appropriate Time) qui est une approximation du 4D-Var incrémental où l’évolution de l’incrément est supposée stationnaire durant la fenêtre d’assimilation, c’est à dire:

$$\mathcal{M}_{0 \rightarrow i}(\mathbf{x}_0^b + \delta \mathbf{x}_0) - \mathcal{M}_{0 \rightarrow i}(\mathbf{x}_0^b) \simeq \delta \mathbf{x}_0$$

Cette formulation de NEMOVAR (décrite dans [Mogensen et al. \(2009\)](#)) sert maintenant de base aux systèmes d’analyse océanique de l’ECMWF et du Met Office.

Afin d’améliorer plus avant ce système j’ai donc proposé la création du projet VODA qui regroupe, en plus de membres de la communauté NEMOVAR, des océanographes du LEGI et du LPO ainsi que des membres de l’équipe système de NEMO. Ce projet qui a débuté en 2009 a plusieurs objectifs, une première phase de développements pour améliorer les solutions présentes dans NEMOVAR puis une série d’application à des échelles différentes (local 1/12°, régional 1/4° et global 1°). VODA a d’ores et déjà permis de mener à bien l’écriture des modèles tangent et adjoint de NEMO (NEMOTAM, [Vidard et al. \(2010\)](#)) et de les fournir à la communauté des utilisateurs de NEMO en lien avec l’équipe gérant les développements et la distribution de ce dernier. Ce projet a aussi permis des avancées importantes dans le développement des matrices de covariance d’erreurs d’ébauche (formulation implicite de \mathbf{B} , [Mirouze and Weaver \(2010\)](#)) et d’observation (\mathbf{R} avec des corrélations entre observations le long des traces altimétriques), ainsi que dans la préparation de futures méthodes multi-échelles (implémentation de l’opérateur de simplification et de son inverse généralisé, [Vidard \(2010\)](#)). Un volet de recherche est également mené sur les stratégies de minimisation, comme par exemple l’assimilation progressive ([Luong et al. \(1998\)](#)) ou quasi statique chez [Pires et al. \(1996\)](#)), qui a été utilisé avec succès dans le cadre d’assimilation d’images par [Titau et al. \(2010\)](#) ou dans un modèle réaliste de l’atlantique Nord dans le cadre de VODA dans [Ferron \(2011\)](#).

1.2 Opérateurs de covariances d’erreurs

Il est impossible de présenter de façon un peu sérieuse les méthodes d’assimilation de données sans s’attarder sur la description des statistiques d’erreur à priori, c’est à dire les matrices \mathbf{B} et \mathbf{R} de l’équation 1.3. Dans cette section on s’intéressera plus au cas de la matrice de covariance d’erreur d’ébauche, considérant, comme on le fait trop souvent que la matrice \mathbf{R} peut être bien représentée par une matrice diagonale. C’est l’approximation qui est généralement faite, bien que, pour de nombreux cas elle est loin d’être valide. C’est le cas même pour des données très utilisées, comme par exemple pour les données altimétries précitées ou les Atmospheric Motion Vectors ([Bormann et al., 2003](#)). Dans ces cas là on augmente artificiellement les variances d’erreur prescrites pour tenter de compenser le manque d’éléments non-diagonaux. Cette approche est loin d’être satisfaisante, comme l’ont montré [Stewart et al. \(2008\)](#) sur un cas simple. Outre les travaux précités, nous sommes en train d’effectuer des investigations sur ces aspects dans le cadre de l’assimilation d’images présentée au chapitre 2.

1.2.1 Le rôle de \mathbf{B}

3D-Var et OI

Les méthodes dites de 3D-Var et d'interpolation optimale (OI) sont des dérivés des méthodes variationnelles et de filtrage respectivement, dans lesquelles le modèle est supposé stationnaire sur la fenêtre d'assimilation et les statistiques d'erreurs n'évoluent pas. Ce sont ces méthodes qui sont actuellement majoritairement utilisées en océanographie opérationnelle. Ces deux algorithmes sont strictement équivalents en théorie et donnent, dans le cas d'un opérateur d'observation linéaire l'état analysé suivant

$$\mathbf{x}^a = \mathbf{x}^b + \mathbf{B}\mathbf{H}^T(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1}(\mathbf{y}^o - \mathbf{H}(\mathbf{x}^b)) \quad (1.16)$$

Afin de bien comprendre les processus en œuvre, une bonne pratique est d'effectuer des expériences mono-observations (on n'assimile qu'une observation unique à la fin de la fenêtre d'assimilation). Dans ce cas, si on suppose de plus que l'observation est située sur un point de grille qui correspond au $k^{\text{ème}}$ élément du vecteur d'état, l'opérateur d'observation devient extrêmement simple.

$$\mathbf{H} = (0 \quad \dots \quad 0 \quad 1 \quad 0 \quad \dots \quad 0)$$

La $k^{\text{ème}}$ colonne est à 1, les autres sont fixées à 0. Ici, $\mathbf{R} = (\sigma_o^2)$ et $\mathbf{H}\mathbf{x}^f = \mathbf{x}_k^f$, et $\mathbf{B}\mathbf{H}^T$ est la $k^{\text{ème}}$ colonne de $\mathbf{B} : (B_{1,k}, B_{2,k}, \dots, B_{n,k})^T$, ainsi

$$\mathbf{x}^a - \mathbf{x}^b = \frac{\mathbf{y}^o - \mathbf{x}_k^f}{\sigma_o^2 + B_{k,k}} \begin{pmatrix} B_{1,k} \\ \dots \\ B_{n,k} \end{pmatrix} \quad (1.17)$$

L'incrément d'analyse est donc proportionnel à une colonne de \mathbf{B} . De par ce fait, le choix de \mathbf{B} est crucial : il va décrire comment cette observation va influencer les points de grille voisins et les autres variables (non observées). On dit classiquement que \mathbf{B} a plusieurs rôles : 1- propagation de l'information : dans les zones peu observées ce sont les structures des covariances d'erreur de l'ébauche prescrites qui vont définir la forme de l'incrément. 2- Lissage de l'information : dans les zones densément observées un lissage est opéré par les corrélations de \mathbf{B} , assurant ainsi que les échelles représentées dans l'incrément sont compatibles avec celles du modèle. 3- Equilibre de l'incrément : grâce aux blocs non diagonaux qui font le lien entre les différentes variables du vecteur d'état.

Effectuer le même exercice avec les équations du 4D-Var donne :

$$\mathbf{x}_i^a - \mathbf{M}_{0 \rightarrow i} \mathbf{x}^b = \frac{\mathbf{y}^o - \mathbf{x}_{i,k}^f}{\sigma^2} \begin{pmatrix} (M_{0 \rightarrow i} \mathbf{B} M_{0 \rightarrow i})_{1,k} \\ \dots \\ (M_{0 \rightarrow i} \mathbf{B} M_{0 \rightarrow i})_{n,k} \end{pmatrix} \quad (1.18)$$

L'incrément au temps i est donc proportionnel à une colonne de $\mathbf{M}\mathbf{B}\mathbf{M}^T$ qui décrit les covariances d'erreur d'ébauche au temps i . L'algorithme 4D-Var fait donc évoluer implicitement les covariances d'erreur au long de la fenêtre d'assimilation. Cette évolution reste cependant implicite et les statistiques d'erreur a posteriori ne sont pas disponibles aisément, contrairement à ce qui se passe pour les dérivés du filtre de Kalman. Pour l'assimilation variationnelle, la matrice \mathbf{B} joue un autre rôle important : elle permet de résoudre ou d'améliorer les aspect sous-déterminés et mal-conditionnés du problème. En effet le terme d'ébauche permet de bien poser le problème de minimisation, même si le nombre d'observations est peu important. De plus la matrice \mathbf{B} est souvent utilisée pour le préconditionnement de la minimisation en effectuant le changement de variable $\mathbf{v} = \mathbf{B}^{-1/2} \delta \mathbf{x}_0$. Il convient de noter que dans ce cas, l'opération

effectivement menée est l'inverse de ce changement de variable, ainsi c'est \mathbf{B} et non son inverse qui est utilisée.

Il est donc clair qu'une bonne représentation de ces statistiques est nécessaire, cependant, si N est la taille du vecteur d'état, \mathbf{B} est de taille N^2 . N étant de l'ordre de 10^7 à 10^{10} pour les applications actuelles en géophysique, cela rend la manipulation de telle matrices inaccessible. Il y a en pratique deux façons de gérer cet aspect. Premièrement on peut décrire cette matrice dans un espace réduit, ce qui est généralement employé dans la plupart des implémentations sous-optimales de filtre de Kalman, comme par exemple le système actuellement opérationnel à Mercator-Océan. L'autre voie, que l'on va développer ici, est de ne jamais former, ni stocker \mathbf{B} , mais de définir un opérateur qui prendra un vecteur \mathbf{v} en entrée et donne $\mathbf{B}\mathbf{v}$ ou $\mathbf{B}^{1/2}\mathbf{v}$ en sortie.

Si ε^b représente l'erreur d'ébauche, alors la matrice \mathbf{B} est définie par

$$\mathbf{B} = \overline{(\varepsilon^b - \bar{\varepsilon}^b)(\varepsilon^b - \bar{\varepsilon}^b)^T} \quad (1.19)$$

, si on note $B_{i,j}$ les coefficients de la matrice \mathbf{B} alors on a $B_{i,j} = \overline{(\varepsilon_i^b - \bar{\varepsilon}_i^b)(\varepsilon_j^b - \bar{\varepsilon}_j^b)}$ Les variances sont alors définies par les éléments diagonaux de $B_{i,i} = \overline{(\varepsilon_i^b - \bar{\varepsilon}_i^b)^2}$ et les corrélations sont les coefficients $C_{i,j}$ tels que $C_{i,j} = \frac{B_{i,j}}{B_{i,i}^{1/2}B_{j,j}^{1/2}}$. Ainsi, si on pose $\Sigma = \text{diag}(B_{i,i}^{1/2})$, on a :

$$\mathbf{B} = \Sigma \mathbf{C} \Sigma \quad (1.20)$$

Σ est la matrice diagonale des écarts types, qui peut être estimée en effectuant des statistiques sur un ensemble suffisamment grand d'états du modèle. Elle ne contient que N éléments non nuls et est donc manipulable.

La matrice \mathbf{C} est sans dimension et représente les corrélations entre les différentes composantes du vecteur d'état, soit entre deux points de grille pour la même quantité (corrélations *monovariées*), soit entre deux différentes quantités (corrélations *multivariées*).

1.2.2 Corrélations monovariées

Dans cette section, on va présenter un moyen de modéliser les corrélations monovariées, c'est à dire représentant les blocs diagonaux de \mathbf{C} , en supposant les corrélations spatiales gaussiennes.

L'approche historique est de définir les corrélations comme une fonction de la distance entre deux points suivant la formule :

$$C_{i,j} = c_1 \exp\left(-\frac{d_{i,j}}{c_2}\right)$$

où c_1 et c_2 sont des facteurs de normalisation et $d_{i,j}$ est la distance entre les points de grille i et j , de sorte que l'influence du point i sur j décroît avec la distance entre ces points. Cette approche a été utilisée dans la partie analyse océanique des système 1 et 2 de prévision saisonnière à l'ECMWF (Anderson et al., 2003) dans un système de type Interpolation Optimale (Eqn: 1.16) avec la limitation d'indépendance niveau par niveau de l'analyse. Pour le développement du système 3 nous n'avions pas la latitude de remplacer ce système archaïque (ce qui a été fait pour le système 4 avec l'adoption de NEMOVAR), j'ai toutefois tenté d'améliorer la spécification de ces corrélations spatiales. La première étape fut de généraliser la formulation à des corrélations 3D. Ce ne fut pas une simple tâche, une analyse indépendante niveau par niveau apportant son lot de simplifications. De plus l'implémentation du système précédent ne facilitait pas son évolution.

Au delà de ces aspects techniques difficiles, la science derrière cette étape est relativement simple. Dans les systèmes 1 et 2, les coefficients de C sont modélisés par

$$C = e^{-\frac{1}{2}\left(\frac{\Delta r}{W_r^*}\right)} \quad (1.21)$$

où Δr représente la distance horizontale entre deux points de grille donnés et W_r^* un facteur d'échelle de corrélation.

Dans le système 3 l'assimilation de la température est effectuée sur tous les niveaux du modèle en même temps, des corrélations verticales ont donc été ajoutées:

$$C = e^{-\frac{1}{2}\left(\frac{\Delta r}{W_r} + \frac{\Delta z}{W_z}\right)} \quad (1.22)$$

où Δz est la différence en profondeur entre deux point donnés. De plus, afin de prévenir l'influence d'une observation de s'étaler au travers des masses d'eau aux caractéristiques vraiment différentes, une dépendance à la densité a été ajoutée. L'équation (1.22) devient :

$$C = e^{-\frac{1}{2}\left(\frac{\Delta r}{W_r} + \frac{\Delta z}{W_z} + \frac{\Delta \rho}{W_\rho}\right)} \quad (1.23)$$

avec $\Delta \rho$, la différence de densité de l'ébauche entre deux points donnés; rendant ainsi les statistiques d'erreur dépendantes de l'état actuel ("error of the day").

La figure 1.3 montre l'incrément d'assimilation venant de 3 observations de température isolées le long de l'équateur pour les formulations 1.22 (haut) et 1.23 (bas). Les contours noirs représentent les isothermes de l'ébauche. On peut noter les différences de comportement décrites ci avant : la diffusion de l'information n'est pas uniforme et dépend de la densité; par exemple, à 200E/210m où les masses d'eau sont bien stratifiées, l'incrément à la forme d'un ellipsoïde alors qu'à 120W/100m sa forme est plus sophistiquée.

Une autre caractéristique du système 3 est l'assimilation de salinité. L'assimilation des données de salinité dans les modèles de circulation générale pour l'océan et le climat est un problème d'importance. Historiquement il y avait peu de données de salinité, mais le récent développement du réseau Argo a radicalement changé la situation. Dans le système 3, les données de salinité furent assimilées après la température (une assimilation simultanée aurait été préférable, mais cela s'avérait difficile techniquement). Dans Haines et al. (2006) nous avons proposé un schéma d'assimilation de salinité basé sur la conservation des propriétés des masses d'eau, permettant de plus grandes échelles de corrélation. Ceci était légitimé par la figure 1.4 qui illustre les échelles de corrélation de la salinité sur l'horizontal (gauche) et sur les isothermes (droite) pour deux points du globe. Ce schéma a été amélioré utilisant le même type d'approche que 1.23 et a été une composante du système 3. Il n'y a malheureusement pas d'étude détaillée et séparée de l'apport de ces améliorations mais une comparaison plus générale avec le système précédent est disponible dans Balmaseda et al. (2007e)

Des méthodes relativement frustrées et peu coûteuses, comme celles utilisées ci-avant permettent des subtilités qui peuvent améliorer le résultat à court terme. Sur le long terme il est sans doute préférable d'investir dans des méthodes plus sophistiquées et plus solides d'un point de vue théorique. On vient de voir que le Système 3 permet de décrire des covariances d'erreur dépendantes de l'état courant (l'erreur du jour) et de rang complet. Parvenir au même résultat avec les approches classiques pour les méthodes variationnelles n'est pas simple, même si en théorie le 4Dvar le fait implicitement. Les filtres le font de façon naturelle, mais ne conservent pas le rang complet des matrices. Il y a cependant plusieurs pistes pour y parvenir, parmi celles ci on peut citer Yaremchuk et al. (2011) qui proposent de combiner un opérateur de rang complet qui ne dépend pas de l'état courant avec une matrice de rang réduit représentant les statistiques

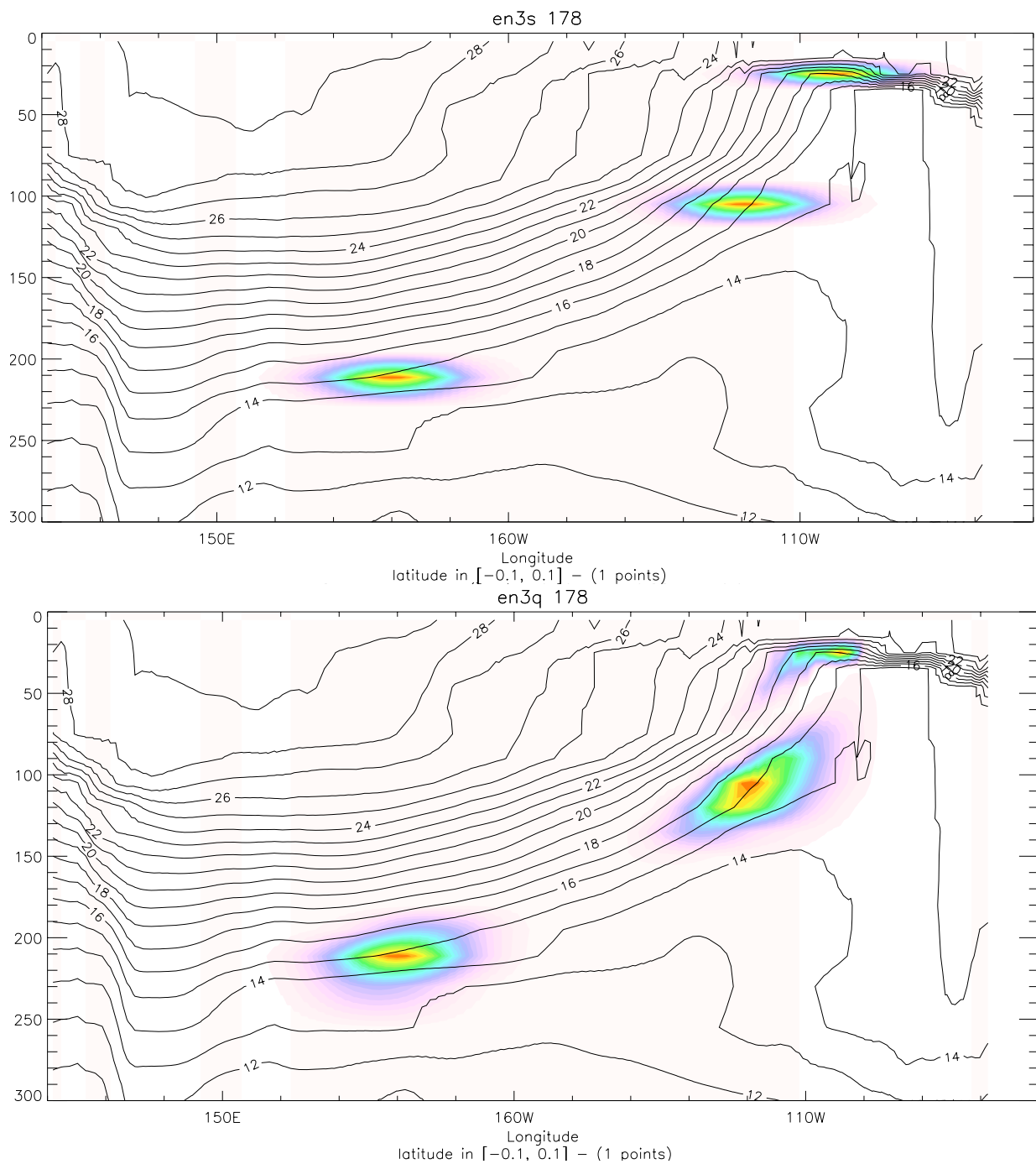


FIGURE 1.3: Impact de trois observations individuelles pour la formulation de **B** traditionnelle (haut) et de la formulation “isopycnale” (bas)

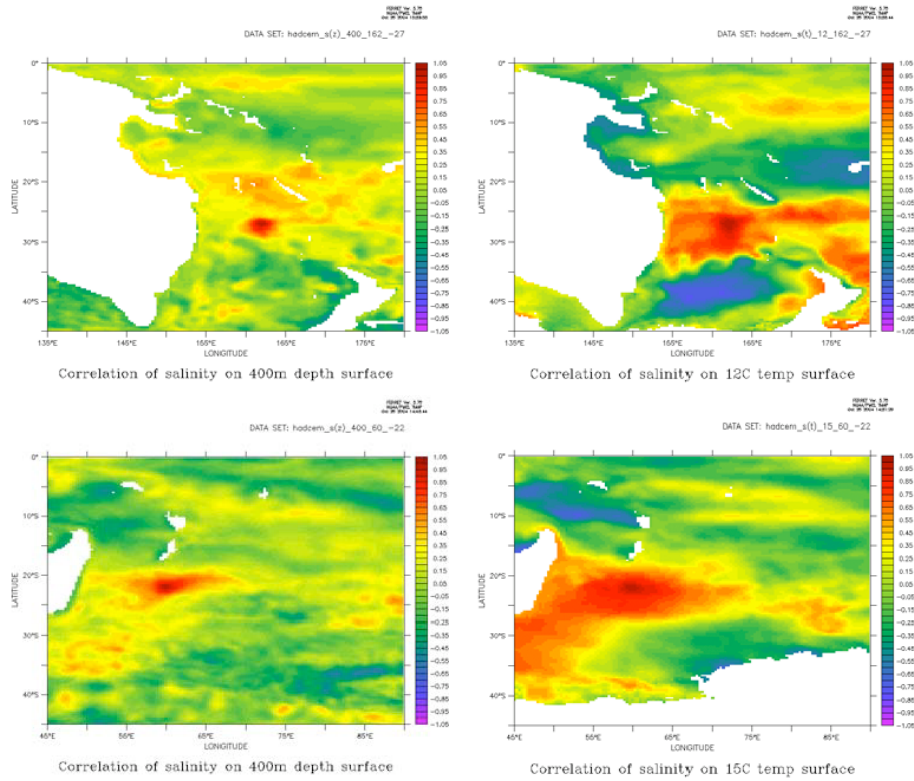


FIGURE 1.4: Cartes de corrélation pour $S(z)$ (gauche) et $S(T)$ droite à 400m issues de 10 ans de simulation haute résolution du modèle HadCEM, pour des points situés à 162E, 27S au large de l’Australie dans le sud Pacifique (haut) et 60E, 22S à l’est de Madagascar dans l’océan Indien (bas)

du jour. De façon plus élégante, mais aussi plus difficile à mettre en œuvre, [Weaver and Mirouze \(2012\)](#) proposent une extension de la méthode maintenant classique basée sur une équation de diffusion généralisée, cette dernière étant présentée dans la section suivante.

Opérateur de diffusion généralisée

L’approche naïve présentée précédemment ne permet pas de gérer correctement les obstacles comme par exemple les côtes en océanographie (l’influence d’une observation peut aisément traverser l’isthme de Panama sans s’acquitter du péage) et ainsi oblige à ignorer les observations trop près des côtes. La méthode dite de l’opérateur de diffusion généralisée améliore l’idée précédente en déformant la structure gaussienne qui va suivre la côte ou le flot. Comme dit précédemment, on a généralement besoin du produit matrice-vecteur plutôt que de la matrice elle-même ou de son inverse. Sous certaines hypothèses (homogénéité et isotropie de la fonction de corrélation) on peut écrire ce produit matrice vecteur sous forme d’un produit de convolution.

Le calcul de ce produit peut-être effectué au moyen de l’intégration d’une équation de diffusion (dont la solution est justement la convolution d’un vecteur initial par un noyau gaussien).

$$\frac{\partial \eta}{\partial t} - \kappa \frac{\partial^2 \eta}{\partial z^2} = 0 \quad (1.24)$$

avec κ un coefficient de diffusion. La solution à la fin d'un intervalle de pseudo-temps $0 \leq t \leq T$ est donné par l'équation intégrale :

$$\eta(z, T) = \frac{1}{\sqrt{4\pi\kappa T}} \int_{z'} e^{-(z-z')^2/4\kappa T} \eta(z', 0) dz' \quad (1.25)$$

$\eta(z, T)$ est ainsi la convolution de $\eta(z, 0)$ avec une fonction de covariance gaussienne où le produit $2\kappa T$ peut être interprété comme le carré de l'échelle de corrélation de la fonction gaussienne et le coefficient constant $(\sqrt{4\pi\kappa T})^{-1}$ comme la variance. Pour revenir à une fonction de corrélation on post multiplie $\eta(z, T)$ par $\sqrt{4\pi\kappa T}$.

Sur la sphère, le calcul de la normalisation est plus compliqué et n'est pas disponible analytiquement. Ainsi on doit calculer une matrice de normalisation qui va convertir la sortie du filtre laplacien \mathbf{L} (équation 1.24) en une multiplication par une matrice de corrélation (avec des 1 sur la diagonale). Celle ci peut être calculée en appliquant successivement \mathbf{L} à tous les vecteurs de la base euclidienne $\mathbf{e}_i = (0, \dots, 0, 1, 0, \dots, 0)$ ou approximée par une méthode aléatoire. Ce calcul est très coûteux ce qui limite la latitude pour faire évoluer les caractéristiques de cet opérateur.

On obtient alors l'opérateur \mathbf{C} de corrélation sur la grille en composant

$$\mathbf{C} = \mathbf{\Lambda} \mathbf{L} \mathbf{\Lambda} \quad (1.26)$$

Cette approche est bien présentée dans [Weaver and Courtier \(2001\)](#) et est actuellement utilisée dans plusieurs systèmes, dont NEMOVAR. En définissant κ comme un tenseur dépendant de l'état courant du système plutôt que comme un coefficient unique, on peut étendre cette méthode afin de modéliser l'action d'une matrice de covariance anisotrope et dépendante de l'“erreur du jour”. Cependant l'implémentation pratique n'est pas triviale et est actuellement le sujet d'une collaboration au sein de VODA / NEMOVAR.

Jusqu'ici seules les covariances spatiales sont prises en compte et les covariances croisées entre les différentes variables du vecteur d'état (les blocs off-diagonaux de \mathbf{B}) ne sont pas encore modélisés, c'est le but de la section suivante.

1.2.3 Corrélations multivariées

Afin de compléter notre opérateur et obtenir une matrice complète (multivariée) de covariance d'erreur d'ébauche, on peu écrire suivant [Derber and Boutier \(1999\)](#)

$$\mathbf{B} = \mathbf{K}_b \hat{\mathbf{B}} \mathbf{K}_b^T \quad (1.27)$$

où $\hat{\mathbf{B}}$ est l'opérateur de covariance monovarié tel que définit dans la section précédente et \mathbf{K} est l'opérateur d'équilibre qui permet d'imposer au résultat de satisfaire à des propriétés physiques connues. Par exemple, dans [Vidard et al. \(2004\)](#), pour un modèle aux équations de Saint-Venant, cet opérateur impose au résultat de satisfaire à l'équilibre géostrophique, l'opérateur monovarié reposant sur l'équation de diffusion généralisée. La figure 1.5 montre, en utilisant ce même système et en assimilant un sous échantillon de la hauteur d'eau, l'évolution de l'erreur RMS pour l'utilisation de l'opérateur monovarié seul (à gauche) et de l'opérateur multiplicité (à droite). Ceci illustre bien l'importance de ces opérateurs d'équilibre pour éviter la génération d'oscillations non voulues (ondes de gravité).

Pour des applications réalistes il n'est pas toujours aisé de trouver des propriétés d'équilibre générales. Pour l'océan les propriétés généralement utilisées sont la conservation des propriétés des masses d'eau et l'équilibre géostrophique ([Weaver et al., 2005](#)), mais elle ne sont pas valides partout, par exemple elle ne le sont pas dans la couche de mélange et à l'équateur respectivement, leur mise en place n'est donc pas

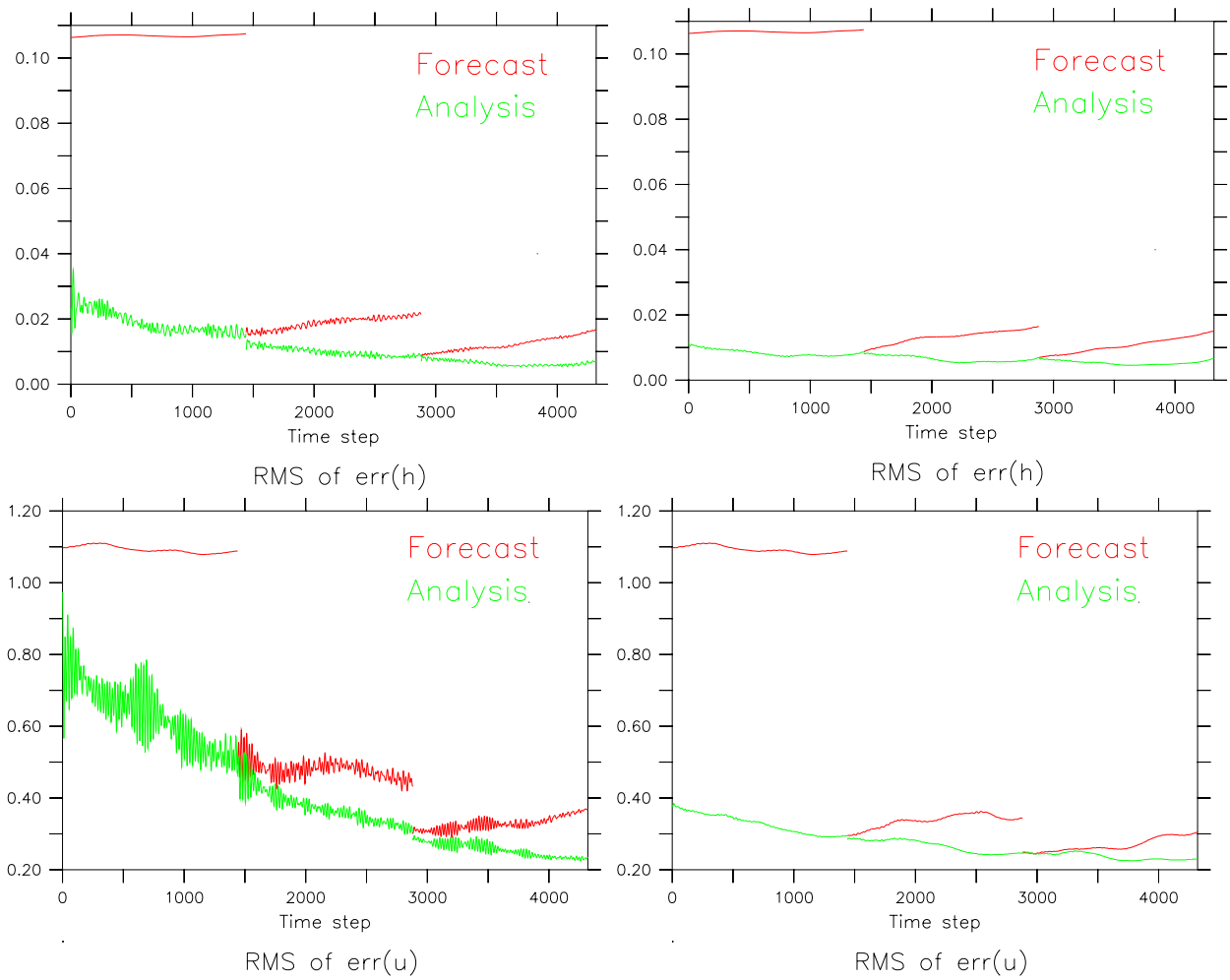


FIGURE 1.5: Courbes d'erreur RMS en fonction du temps pour l'ébauche (rouge) et l'analyse (vert) de trois cycles successifs de 4Dvar dans modèle aux équations de Saint Venant, pour la hauteur d'eau (haut) et la vitesse (bas) avec un opérateur \mathbf{B} monovarié (gauche) et multivarié (droite).

toujours aisée. Cependant, avec le 4D-Var, si il y a assez d'observations et si la fenêtre d'assimilation est suffisamment longue, l'état analysé et implicitement contraint à l'équilibre, par la présence du modèle et de l'adjoint dans le système d'optimalité ; le besoin d'un tel opérateur sophistiqué en est donc réduit. C'est d'ailleurs la tendance actuelle en prévision numérique du temps où le nombre d'observation a augmenté de façon importante, ainsi le besoin de l'opérateur \mathbf{K} se fait moins sentir (Y. Trémolet, communication personnelle). Pour le pauvrement observé océan, cet aspect reste important.

1.3 Le contrôle d'erreur modèle en assimilation variationnelle

Dans les méthodes classiquement utilisées en assimilation de données variationnelles telles que décrites au début de ce chapitre le terme de contrôle est la condition initiale du système. Ce faisant on suppose implicitement que le modèle est parfait. De toute évidence cette hypothèse n'est pas réaliste. Au cours de ma thèse, j'ai proposé plusieurs pistes pour prendre en compte l'erreur modèle. Celles ci considèrent l'erreur de modèle comme un terme supplémentaire que l'on peut contrôler.

Le modèle s'écrit alors

$$\begin{cases} \frac{d\mathbf{x}}{dt} &= \mathcal{M}(\mathbf{x}) + \eta(t), \quad t \in [0, T], \\ \mathbf{x}(0) &= \mathbf{x}_0, \end{cases} \quad (1.28)$$

et la fonction coût à minimiser devient

$$J(\mathbf{x}_0, \eta) = \frac{1}{2} \|\mathbf{x}_0 - \mathbf{x}^b\|_{\mathcal{X}}^2 + \frac{1}{2} \int_0^T \|\mathbf{y}(t) - \mathcal{H}[\mathbf{x}(\mathbf{x}_0, t)]\|_{\mathcal{O}}^2 dt + \frac{1}{2} \int_0^T \|\eta(t)\|_{\mathcal{E}}^2 dt, \quad (1.29)$$

Cette approche n'est pas réalisable en pratique pour des problèmes de grande taille : la taille du vecteur de contrôle devient immense (taille du vecteur d'état \times nombre de pas de temps). Pour circonvier à ce problème plusieurs approches ont été proposées. Une première possibilité est de modéliser l'évolution du terme correctif d'erreur et de contrôler que sa condition initiale (Griffith and Nichols, 2000; Vidard et al., 2004). On suppose que l'on connaît un modèle d'évolution de l'erreur:

$$\begin{cases} \frac{d\eta}{dt} &= \mathcal{N}(\eta(t), \mathbf{x}(t)) + \epsilon(t), \quad t \in [0, T], \\ \eta(0) &= \eta_0, \end{cases} \quad (1.30)$$

où ϵ est un bruit stochastique non biaisé, qui peut éventuellement être contrôlé (mais dans la suite on le négligera), et la fonction coût devient

$$J(\mathbf{x}_0, \eta_0) = \frac{1}{2} \|\mathbf{x}_0 - \mathbf{x}^b\|_{\mathcal{X}}^2 + \frac{1}{2} \int_0^T \|\mathbf{y}(t) - \mathcal{H}[\mathbf{x}(\mathbf{x}_0, t)]\|_{\mathcal{O}}^2 dt + \frac{1}{2} \|\eta_0 - \eta^b\|_{\mathcal{E}}^2 dt. \quad (1.31)$$

Cette approche a été utilisée avec succès dans des conditions pré-opérationnelles (modèle réaliste, données réelles) pour l'océan (Vidard, 2001) et pour l'atmosphère (Trémolet, 2007c) en considérant η comme étant constant sur une fenêtre d'assimilation, on parle alors de contrôle de l'erreur *systématique*. Vidard et al. (2004) a même montré sur un cas simple que l'estimation de cette erreur pouvait éventuellement être utilisée pour corriger la prévision subséquente.

Une autre option, pour réduire la taille du vecteur de contrôle due aux termes de correction d'erreur est de les projeter sur une base réduite comme proposé dans certains de mes articles (Vidard et al., 2000, 2003,

2004) et dans les travaux de S. Durbiano (2001). Dans ce cas l'équation 1.28 s'écrit

$$\begin{cases} \frac{d\mathbf{x}}{dt} &= \mathcal{M}(\mathbf{x}) + \mathbf{T}\eta_r(t), \quad t \in [0, T], \\ \mathbf{x}(0) &= \mathbf{x}_0, \end{cases} \quad (1.32)$$

avec η_r dans un espace de petite dimension.

On peut même envisager une approche regroupant les deux stratégies citées ci dessus, comme illustré dans Vidard et al. (2000) sur un cas très simple.

Plus récemment Trémolet (2007a) a proposé de retourner à la formulation générale dans laquelle c'est la trajectoire du modèle toute entière qui est contrôlée. Ce qui donne pour un modèle discrétisé en temps

$$J(\mathbf{x}) = \frac{1}{2} \sum_{i=1}^I \| \mathbf{y}_i - \mathcal{H}_i[\mathbf{x}_i] \|_{\mathcal{O}}^2 + \frac{1}{2} \| \mathbf{x}_0 - \mathbf{x}^b \|_{\mathcal{X}}^2 + \frac{1}{2} \sum_{i=1}^I \| \mathcal{M}_{t_{i-1} \rightarrow t_i}(\mathbf{x}_{t_{i-1}}) - \mathbf{x}_{t_i} \|_{\mathcal{E}}^2 \quad (1.33)$$

On se retrouve dans le même problème que 1.29, à savoir une taille de vecteur de contrôle ingérable. Trémolet (2007a) propose alors de ne contrôler qu'un sous-ensemble réduit d'instantanés de cette trajectoire en pénalisant l'écart à l'intégration du modèle depuis l'instantané précédent. En limitant le nombre d'instantané, cet algorithme a un coup raisonnable, de plus le calcul du terme d'écart aux observations, qui est le plus coûteux car requérant l'intégration rétrograde de l'adjoint, ce calcul, donc, est indépendant sur chaque sous-fenêtre et peut donc être effectué en parallèle. Cette approche ressemble à ce que certains appellent en automatique le *Moving Horizon Estimator* avec *direct multiple shooting* décrit dans Kraus et al. (2006) et bibliographie. A ma connaissance, cette approche n'a pas été utilisée sur des modèles réalistes. Y. Trémolet et moi même avons effectué son implémentation dans le système OPAVAR (ancêtre de NEMOVAR) lors de sa visite dans notre équipe en 2006, malheureusement nous n'avons pas eu le temps d'obtenir de ces travaux des résultats publiables avant de devoir retourner à nos occupations respectives. Ce n'est que partie remise, B. Lemieux et moi même travaillons à son installation dans NEMOVAR dans le cadre du projet VODA (Lemieux and Vidard, 2011) tandis que Y. Trémolet participe au développement de OOPS (An Object Oriented Framework for Coupling Data Assimilation Algorithms to Models) dont une des motivations est d'apporter la souplesse nécessaire pour pouvoir implémenter ce type de méthode dans le cadre du système intégré de prévision de l'ECMWF (IFS).

Le contrôle de l'erreur modèle ou "4DVar à contrainte faible" (Dans le 4D-Var classique le modèle est une contrainte forte de la minimisation) présente certains avantages, le premier d'entre eux est qu'il permet l'utilisation de fenêtres d'assimilation plus grandes (Vidard, 2001). Ce résultat est assez intuitif : si le modèle dérive il devient difficile d'ajuster la condition initiale de façon cohérente afin de s'approcher des observations dans leur ensemble. Un autre raison, est que la quasi totalité des systèmes utilisent une méthode de type Gauss-Newton (ou 4D-Var incrémental, *i.e.* la minimisation de Eq1.13) qui effectue la minimisation par approximation successive utilisant la TLA (Equation 1.9). Si la TLA est trop grossière cet algorithme ne converge pas, or celle ci est directement dépendante de la longueur de la fenêtre d'assimilation. Dans le cas du 4DVar contrainte faible, la TLA ne doit être fine que sur l'intervalle entre deux corrections de la trajectoire. L'allongement des fenêtre d'assimilation rend moins important le besoin de bien spécifier l'opérateur d'équilibre (équation 1.27) dans l'opérateur de covariance d'erreur d'ébauche. Un autre aspect est qu'il permet une plus grande parallélisation, comme mentionné plus haut. Cet aspect est encore renforcé par l'utilisation de méthodes dites de "point-selle" comme proposées par Lagarde (2000) et redécouvertes par Fisher and Auvinen (2012) où une méthode 4DVar et une méthode duale (4D-PSAS) sont combinées pour augmenter le parallélisme.

1.4 Quelques pistes pour l'avenir

Comme on l'a vu lors de ce chapitre, les méthodes d'assimilation ont connu un très fort développement ces trente dernières années, elles ont maintenant atteint une certaine maturité et sont utilisées de façon opérationnelle dans de nombreux domaines. Ces méthodes ont permis, notamment dans les systèmes de prévision météorologiques, malgré des non-linéarités importantes, d'utiliser une grande variété et une grande quantité d'observations pour générer des états analysés convenablement équilibrés et de meilleure qualité, et ainsi d'améliorer grandement les performances desdits systèmes. Cependant l'histoire ne s'arrête pas là et l'effort de recherche va se poursuivre tellement les enjeux sont importants. Sans prétendre être exhaustif, on peut tout de même, en guise de conclusion à ce chapitre sur les méthodes, tenter de dégager quelques grands défis pour les années à venir dans ce domaine.

1.4.1 Evolution des méthodes et adaptation à l'évolution des moyens de calcul

Après avoir vu l'avènement des architectures vectorielles puis sa marginalisation, la "mode" actuelle est au massivement parallèle et ça le restera à moyen et sans doute long terme. Ceci impose de penser les évolutions des méthodes à travers le prisme (pour ne pas dire la dictature) de la "scalabilité". En effet à qualité égale, les méthodes qui équilibrent mieux la charge de calcul seront plébiscitées. En l'état actuel, les méthodes variationnelles utilisées en PNT tournent efficacement sur plusieurs milliers de processeurs. On dispose déjà de calculateurs à plusieurs dizaines, voire centaines de milliers d'unités de calcul, maintenir de bonnes performances sur ce type de calculateur ne va pas être trivial. Au regard de cela, les méthodes d'ensemble ou particulières sont évidemment bien placées, il suffit d'augmenter la taille de l'ensemble. En ce qui concerne les méthodes variationnelles plusieurs pistes semblent se dégager, le 4DVar à contrainte faible tel que décrit à la fin de la section 1.3 fournit un élément de réponse permettant un niveau de parallélisme plus important et de plus grandes fenêtres d'assimilation. Ce dernier aspect a d'ailleurs un effet double : il autorise de prendre en compte un nombre plus important d'observation pour une seule analyse, il permet aussi de réduire la sensibilité à la condition initiale. De plus, le nombre et la variété grandissante des données obtenues grâce à l'important développement des réseaux d'observation, s'il augmente de façon importante le coût de l'EnKF, a de nombreux impacts sur la formulation variationnelle et peut faire changer les priorités : l'erreur sur la condition initiale et les incréments d'analyse sont plus petits, les erreurs sont de plus petite échelle et moins structurées, et surtout l'importance du terme d'ébauche de la fonction coût et des contraintes d'équilibre associée diminuent. On pourrait même imaginer que dans le futur, le problème de minimisation du terme d'écart aux observations seul soit bien posé (tout ceci est surtout vrai pour la météorologie, pour l'océanographie, on en est encore loin).

1.4.2 Assimilation pour les systèmes de modélisation

Dans de nombreux domaines se développent ce qu'il convient d'appeler des systèmes de modélisation, où non plus une seule physique ou une seule échelle est représentée et son évolution décrites, mais plusieurs milieux ou phénomènes interagissent. Ces couplages peuvent être de plusieurs types : avec un autre milieu (pour ce qui concerne l'océan, on peut mentionner le couplage océan-atmosphère, mais aussi océan-biogéochimie), d'un même milieu mais avec des échelles et des approximations différentes (couplage d'un modèle hauturier hydrostatique avec un modèle côtier non hydrostatique par exemple), des modèles emboîtés (un zoom haute résolution dans une région d'intérêt emboîté dans un modèle global à basse résolution), etc. Dans la plupart des systèmes actuels, les composantes de ces systèmes de modélisation sont initialisés par assimilation de données de façon indépendante. Cette technique ad-hoc permet d'utiliser indépendamment

les systèmes d'assimilation développés pour chaque modèle, sans avoir à reconstruire un nouveau système. Par contre, elle est à l'évidence sous-optimale. Dans le cas du couplage océan-atmosphère, par exemple, du fait de l'incohérence relative des forçages provenant de réanalyse et des flux engendrés par un modèle (d'océan ou d'atmosphère), le passage du mode forcé de l'assimilation au mode couplé de la prévision induit un choc dans l'évolution du système (McPhaden et al., 1998). Ce choc peut entraîner, au début de la période de prévision, la création d'ondes inexistantes dans la réalité.

Le développement de méthodes d'assimilation de données performantes pour les modèles couplés est un point crucial : il s'agit maintenant de franchir une étape décisive dans les techniques utilisées, tant probablement dans l'amélioration apportée en terme de qualité des résultats, que dans les problèmes de coût numérique et de complexité numérique supplémentaires que cela va soulever. En effet tenter de rassembler des informations très hétérogènes (coupler des modèles hétérogènes en échelles, multi-fluide, multi-phase et des données de toutes sortes) ne se fera pas aisément. De ce fait, développer méthodes mathématiques et algorithmes efficaces, en insistant sur la parallélisation de ceux-ci, seront des points essentiels.

1.4.3 Évolution de la formulation des statistiques d'erreurs

Une étape importante, à mon sens, est la meilleure prise en compte des erreurs d'observations corrélées entre elles. En effet, avec l'omniprésence des observations satellites avec leur multiples canaux, l'hypothèse de non corrélation de ces erreurs est mise à mal, et négliger ces corrélations croisées peut dégrader l'analyse si les observations sont trop denses (Liu and Rabier, 2006). La solution la plus couramment employée actuellement consiste à sous-échantillonner ces observations et à augmenter artificiellement la variance d'erreur spécifiée correspondante. De nombreuses études ont mis en évidence la présence de ces corrélations croisées, comme par exemple les travaux de N. Bormann au CEPMMT (Bormann et al., 2003; Bormann and Bauer, 2010; Bormann et al., 2010). Il est cependant difficile de définir une stratégie commune à tous types d'observation de part leur différence de nature, et de plus la taille des vecteurs d'observation pouvant être importante, on se retrouve devant le même problème d'impossibilité de manipulation que dans le cas de la matrice d'erreur d'ébauche. La communauté commence à s'intéresser à ce type de problème, soit en manipulant effectivement ces matrices plus ou moins creuses (Weston, 2011) ou en définissant un opérateur de multiplication par cette matrice (ou son inverse) à l'instar de ce qui est décrit dans la section 1.2.2 (Weaver and Piacentini, 2012)

Même si la matrice (ou l'opérateur) de covariance d'erreur d'ébauche a reçu plus d'attention, il reste encore de nombreuses améliorations à apporter. Comme il a été évoqué lors de la section 1.2, trouver un moyen de modéliser les statistiques d'erreur d'ébauche de manière à ce qu'elles soient dépendante de l'état courant est un enjeu d'importance en assimilation variationnelle. En effet dans les implémentations habituelles de cette approche ces statistiques sont fixées au début de chaque fenêtre d'assimilation et en général toujours aux mêmes valeurs. Or, de tout évidence, ces erreurs d'ébauches sont dépendantes de l'état du système, on n'aura pas les mêmes caractéristiques sur une mer d'huile que lors du passage d'un ouragan. Ceci est possible à mettre en œuvre dans le cadre d'un opérateur basé sur la diffusion généralisée mais pose des problèmes de calcul scientifique non triviaux comme par exemple la résolution de nombreux systèmes linéaires de grande taille coûteux à résoudre.

1.4.4 Méthodes hybrides et nouvelles méthodes

Un lecteur éclairé aura remarqué que j'ai jusqu'ici presque passé sous silence les méthodes hybrides 4DVar / Ensemble qui connaissent pourtant un fort développements actuellement. En effet, un autre moyen d'obtenir la dépendance à l'état courant des statistiques d'erreur initiales dans un environnement variationnel

est de faire évoluer ces statistiques par un filtre ou un lisseur équivalent. Dans [Buehner et al. \(2010a,b\)](#) un résultat important vient soutenir l'intérêt d'une telle démarche : si les deux systèmes (4DVar et EnKF) sont grosso modo équivalents, ils obtiennent des résultats sensiblement meilleurs en utilisant les statistiques d'erreur calculées par le filtre de Kalman d'ensemble dans le 4DVar. On pourrait résumer ce résultat en prétendant que, pour ce système au moins, l'assimilation variationnelle est plus efficace pour répondre à la question posée, mais le filtre d'ensemble la pose mieux. En d'autres termes, un bon mélange serait un partage des tâches dans lequel le filtre ou lisseur de Kalman d'ensemble ou assimilé fait évoluer la matrice de covariance d'ébauche pour le 4DVar qui lui, aurait à charge de minimiser la fonction coût ainsi produite. On peut en effet raisonnablement penser que l'approche variationnelle est mieux qualifiée pour manipuler la non-linéarité de la dynamique et des opérateurs d'observation et de plus permet plus facilement d'imposer des contraintes d'équilibre physique.

L'idée de méthodes hybrides n'est pas récente, il y eut quelques tentatives au début des années 2000, par exemple l'algorithme 3DVar/EnKF introduit par [Hamill and Snyder \(2010\)](#) qui utilise l'ensemble pour propager les statistiques dans un cadre non linéaire mais n'exploite pas les propriétés lissantes du 4D-Var ou [Veersé et al. \(2000\)](#) qui propose un cadre théorique pour un système hybride entre un lisseur de Kalman de rang réduit et un 4D-Var de rang réduit, mais sans proposer d'application.

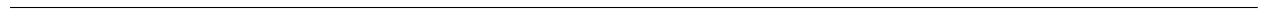
D'autres tentatives de produire des méthodes hybrides tendent plus à recycler (ou importer) des idées d'une approche vers l'autre. Par exemple le 4DEnKF (4D Ensemble Kalman Filter) introduit par [Hunt et al. \(2004\)](#) fournit un moyen pour assimiler les observations passées (et non uniformes) dans un cadre séquentiel, mais n'agit pas intrinsèquement comme un lisseur ni ne se place vraiment dans un cadre 4DVar. Dans l'autre sens on peut citer la méthode VAE (Variational Assimilation Ensemble) opérationnelle à Météo-France et décrite dans [Raynaud et al. \(2011, et bibliographie\)](#) qui effectue six assimilations 4DVar ou 3DFGAT perturbées et découplées en parallèle pour estimer les variances d'erreur, mais n'hybride pas fondamentalement les deux algorithmes.

Plus récemment, plusieurs propositions pour des approches hybrides plus intégrées ont vu le jour. Parmi celles-ci on peut citer la méthode E4DVAR (Ensemble 4DVar) proposée par [Zhang et al. \(2009\)](#) qui effectue un 4DVar et un EnKF en parallèle modifiant la moyenne de l'ensemble en se basant sur le résultat du 4DVar et fournissant au 4D-Var des statistiques d'ébauche basées sur les résultats de l'EnKF.

Une autre approche cohérente est décrite dans les travaux de [Krysta et al. \(2011\)](#) auxquels j'ai été (un peu) associé ([Krysta et al., 2008](#)), et qui sont basés sur les travaux théoriques de [Veersé et al. \(2000\)](#). Il s'agit ici de combiner une 4DVar d'ordre réduit et un lisseur de Kalman d'ordre réduit sur la même base.

Il reste cependant encore un peu de travail avant que ces méthodes hybrides aient la maturité suffisante (à part pour les plus simples) pour être utilisées dans un cadre opérationnel. L'activité de recherche est importante actuellement sur ce type d'approche et le nombre de variantes commence à foisonner, là encore un ensemble de cas test communs pour évaluer leurs forces et leurs faiblesses respectives serait nécessaire.

Un autre aspect est que les méthodes hybrides cohérentes actuellement proposées se marient mal avec l'approche d'un opérateur \mathbf{B} modélisé comme présenté dans la section 1.2. Dans ce cas un système hybride peut aisément fournir une actualisation des variances (les termes diagonaux) et un peu plus difficilement les longueurs de corrélation. Une possibilité probablement abordable serait d'avoir \mathbf{B} représenté en deux parties, une première en base réduite venant du filtre et une deuxième de rang complète, prescrite et modélisée sous forme d'opérateur. On doit pouvoir aller vers des méthodes plus intégrées, mais il reste à les imaginer.



Chapitre 2

DES OBSERVATIONS RÉCALCITRANTES

Les observations sont des mesures partielles, en temps et en espace et discrètes de quantité physiques continues. De plus les observations généralement utilisées en assimilation de données sont en le plus souvent encore sous échantillonnées et filtrées. Ainsi, l'espace des observations \mathcal{O} est un sous espace de \mathbb{R}^p où p est le nombre d'observation. Les observations \mathbf{y} et le vecteur d'état \mathbf{x} sont liés par l'opérateur d'observation

$$\begin{aligned} \mathcal{H} : \mathcal{X} &\rightarrow \mathcal{O} \\ \mathbf{x}(\mathbf{t}) &\mapsto \mathbf{y}(\mathbf{t}) = \mathcal{H}(\mathbf{x}(\mathbf{t})), \end{aligned} \quad (2.1)$$

qui “projette” l'espace des variables d'état \mathcal{X} vers l'espace des observation \mathcal{O} . Pour l'assimilation de données on définit l'espace des observations en lui associant un produit scalaire prenant en compte les incertitudes sur les mesures et les erreurs de représentativité. Ainsi

$$\|\mathcal{H}(\mathbf{x}) - \mathbf{y}\|_{\mathcal{O}}^2 = (\mathcal{H}(\mathbf{x}) - \mathbf{y})^T \mathbf{R}^{-1} (\mathcal{H}(\mathbf{x}) - \mathbf{y}), \quad (2.2)$$

où \mathbf{R} est la matrice de covariance d'erreur d'observation.

On parle d'observations *directes* quand la quantité observée appartient à l'ensemble des quantités qui sont représentées par la variable d'état du modèle et d'observations *indirectes* sinon. Dans le premier cas, \mathcal{H} peut être défini comme une projection de \mathcal{X} vers \mathcal{O} . Dans le second cas, évaluer $\mathcal{H}(\mathbf{x})$ peut être difficile, tant du point de vue calcul que mathématique du fait des lois physiques complexes qui peuvent lier \mathbf{x} et \mathbf{y} . L'opérateur d'observation doit être suffisamment régulier pour que les dérivées partielles $\frac{\partial \mathcal{H}}{\partial \mathbf{x}}$ qui sont requises pour obtenir le modèle adjoint (1.7) existent. Pour des observations directes, cette propriété est généralement facilement satisfaite, alors que, dans le cas des observations indirectes, cela peut être plus délicat.

Par exemple, dans le cas d'un modèle d'océan dont les variables d'état sont les vitesses de courant, la hauteur de la surface, la température et la salinité, des observation directes pourraient être des mesures in situ de température et salinité dans des endroits précis du domaine, alors que des observations indirectes proviendront de mesures satellitaires micro-ondes qui sont liées à la salinité de surface. Dans ce cas \mathcal{H} représentera la relation entre les radiances micro-onde et la salinité.

La partie observation est souvent laissée de côté lorsqu'on dessine une méthode d'assimilation de données. La mise en pratique dans un cadre réaliste ou pis encore opérationnel montre pourtant qu'il s'agit d'un élément crucial et délicat. En effet, Lorsqu'il s'agit d'assimiler de vraies observations, une partie des algorithmes d'assimilation qui semblait aisée : le calcul de $\|\mathcal{H}(\mathbf{x}) - \mathbf{y}\|_{\mathcal{O}}^2$ peut s'avérer difficile. Le premier problème étant bien entendu la possibilité d'observations farfelues (une température de 30° dans l'océan

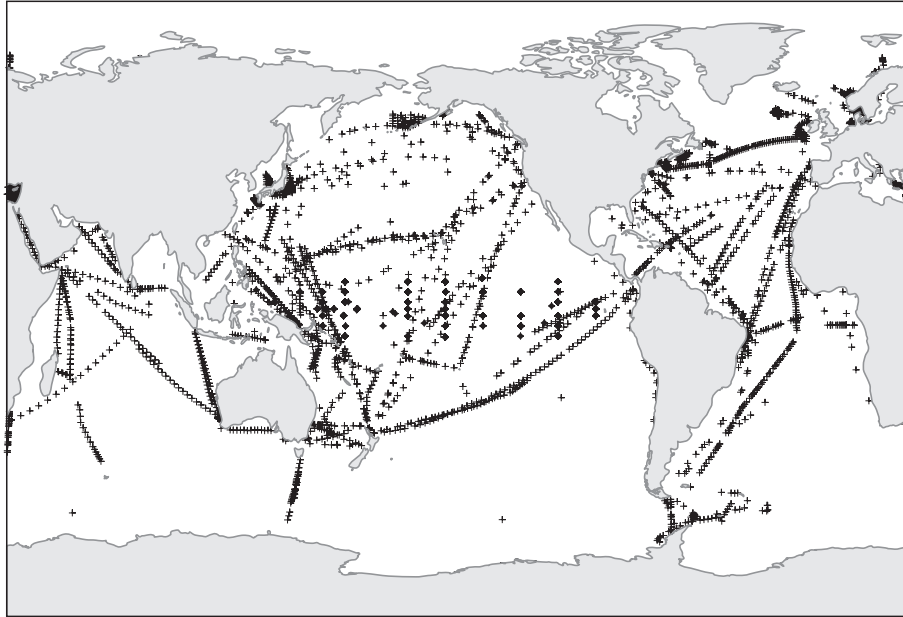


FIGURE 2.1: Observations en température pour 10 jours de mars 1993. Les losanges sont des bouées ancrées et les croix des sondes XBT

Arctique, une ligne XBT naviguant à 300 km/h, un flotteur au milieu du Sahara, ...) ou douteuses (un instrument en désaccord avec les mesures équivalentes à proximité). Celles-ci peuvent être écartées au moyen d'un contrôle de qualité *a priori* (Ingleby and Huddleston (2007)) mais au prix d'une ingénierie complexe.

Plus délicat est le problème engendré par la disparité de la distribution spatiale des observations (cf Fig 2.1) qui peut engendrer une déstabilisation du système d'analyse. Paradoxalement, c'est plutôt une sur-observation d'un domaine qui risque le plus de poser problème. Ceci vient principalement de la mauvaise spécification des erreurs d'observations et principalement de l'hypothèse de non corrélation des erreurs des diverses mesures couramment admise pour des raisons évidentes de simplification (\mathbf{R} est diagonale dans ce cas là). Si on peut discuter de la validité de cette hypothèse pour les erreurs de mesures elle mêmes, elle est de toute évidence invalide pour les erreurs dites de représentativité (phénomènes observables mais non représentés par le modèle en raison d'erreurs de discrétisation ou de physique manquante). Les solutions le plus communément employées sont alors le "thinning" (sous-échantillonnage des observations) ou le "superobbing" (moyenne pondérée d'un groupe d'observations pour créer une "super" observation qui va les remplacer, (Smith et al., 1991)) ou pis encore, la diminution globale du poids des observations, mais ces solutions s'apparentent plus à du bricolage et un travail plus approfondi sur les spécification des statistiques serait préférable, même pour des observations ponctuelles.

L'évolution dans le temps du réseau d'observation peut également causer soucis au malheureux concepteur d'un système d'assimilation opérationnel. On peut voir, par exemple, sur les Fig 2.1 et Fig 2.2 l'évolution du réseau *in situ* océanographique entre 1993 et 2011 avec l'apparition des flotteurs Argo et la quasi disparition des sondes XBT lancées par les navires d'opportunités. Historiquement, l'océan a été assez peu observé (surtout dans l'hémisphère sud), et principalement par des mesures de température, même si les différents réseaux d'observations pouvaient avoir un impact différent (Vidard et al. (2007); Balmaseda et al. (2007a)). Cet aspect monovarié des observations rendait la mise en œuvre des méthodes d'assimilation facilitée (pas forcément la qualité de l'analyse), la donne a changée avec l'apparition des données satellitaires au début des

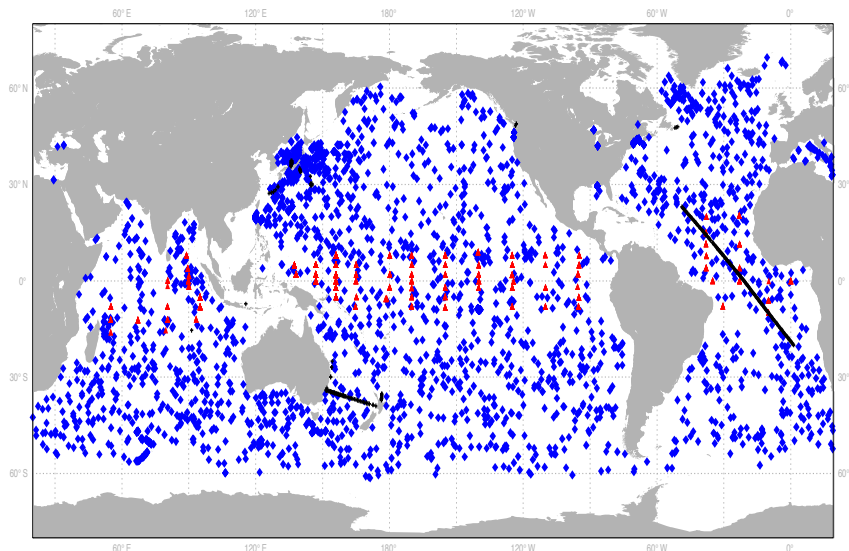


FIGURE 2.2: Observations en température pour 10 jours en juillet 2011. Les triangles rouges sont les bouées ancrées, les croix noires sont les sondes XBT et les losanges bleus sont les flotteurs Argo

années 1990 et de l'ajout presque systématique de données de salinité au début des années 2000. L'introduction brutale de certaines données nombreuses comme par exemple les observations altimétriques satellitaires en 1993 (Fig 2.3) ou les bouées ancrées TRITON au début des années 2000 peuvent, si l'on n'y prend garde, déstabiliser le système et même une introduction graduelle comme celle des flotteurs Argo peut rendre certaines hypothèses sur les statistiques d'erreur caduques. Ces variations doivent particulièrement être prise en compte dans le cas d'une longue réanalyse où la cohérence temporelle est importante (Balmaseda et al., 2008, 2007c).

Le contrôle de qualité des observations est un travail délicat d'ingénierie qui peut parfois s'apparenter à du bricolage, il existe cependant des approches plus solides (et complémentaires) basées sur l'exclusion ou la réduction de l'influence des observations ne satisfaisant pas aux hypothèses statistiques présupposées (Ingleby and Lorenc, 1993; Andersson and Järvinen, 1999)). Même si l'on fait fi des problèmes d'observation trop ou pas assez nombreuses ou de qualité déplorable, on peut se retrouver confronté à deux types de problèmes plus fondamentaux : 1- le lien entre les variables d'état du modèle peut-être très indirect et 2- les erreurs associées à certains types d'observations sont mal connues, difficiles à spécifier et/ou ne satisfaisant pas aux hypothèses de non corrélation.

Pour illustrer une partie des difficultés qui ont été évoquées dans ce préambule, ce chapitre se divise en trois sections qui présentent deux types de données particulières sur lesquelles j'ai été amené à travailler : - les données altimétriques, cette section résume les difficultés inhérente à ce type de données qui est déjà largement utilisé et dont l'utilisation semble pourtant directe. Ces considérations sont principalement issus des travaux effectués avec Magdalena Balmaseda et David Anderson dans le cadre du développement du système III de prévision saisonnière de l'ECMWF et résumés dans les articles Vidard et al. (2009a); Balmaseda et al. (2008). - Les données de type image. Cette section tire parti des travaux effectués dans le cadre des travaux de doctorat d'Innocent Souopgui et Vincent Chabot et de post-doctorat d'Olivier Titaud (Souopgui (2010); Titaud et al. (2010)) que j'ai eu le plaisir de co-encadrer.

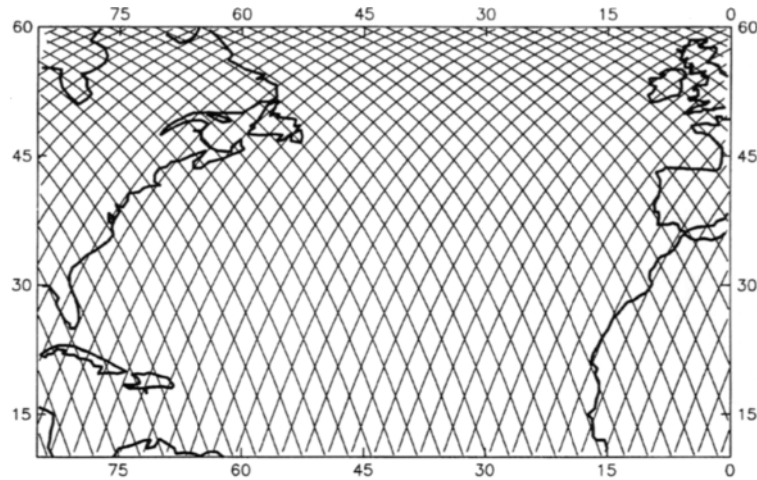


FIGURE 2.3: Data coverage altimetry in the North Atlantic, 10 days of Topex/Poseidon. (source [Fu and Cazenave \(2001\)](#))

2.1 Données altimétriques

2.1.1 Topographie Dynamique Moyenne

Les mesures altimétriques par satellites donne une information sur la hauteur de la surface de la mer (SSH) relative à un ellipsoïde de référence. Cet ellipsoïde n'a pas de signification physique particulière, une mesure plus adaptée serait par rapport au géoïde terrestre (la surface d'équi gravité) mais jusqu'à récemment les estimations du géoïde étaient trop approximatives pour être utilisées dans ce contexte. La hauteur de la surface de la mer relative au géoïde est généralement appelée hauteur dynamique (η , c'est l'influence de l'état actuel local de l'océan) et sa moyenne temporelle la hauteur dynamique moyenne (MDT). Pour éviter le besoin de géoïde, les mesures altimétriques étaient données en anomalies de hauteur de mer (SLA), anomalie relative à une hauteur moyenne (MSSH) mesurée sur période de temps de plusieurs années (tout ceci est schématisé dans la figure 2.4).

Les modèles possèdent généralement maintenant une surface libre η . Dans le monde du modèle, le géoïde est une sphère, de sorte que la SSH et η (et donc la MSSH et la MDT) sont équivalentes. Afin de comparer le modèle et les observations, on a besoin, soit de transformer η du modèle en SLA, soit de ramener la SLA observer à une hauteur dynamique η^{obs} . Dans le premier cas il suffit de retrancher la MDT (donc la MSSH) du modèle, qui est aisément calculable, à η . Cette MDT peut d'ailleurs contenir des information provenant d'autres observations car elle peut être estimée en moyennant la SSH d'une série temporelle provenant d'une analyse océanique assimilant, par exemple, les données de température et salinité. La deuxième approche requiert la connaissance d'une MDT observée. La première approche est plus facile à mettre en œuvre, mais la seconde est préférable puisqu'elle devrait permettre de corriger également l'état moyen.

Avec le lancement de la mission gravimétrique CHAMP (Challenging Minisatellite Payload), GRACE (Gravity Recovery And Climate Experiment) et le plus haute résolution GOCE (Gravity Field and Steady-State Ocean Circulation Explorer) notre connaissance du géoïde s'est grandement améliorée, menant à des estimations de la MDT plus précises, telles que celles fourni par la NASA ([Tapley et al., 2003](#), Tap03) ou

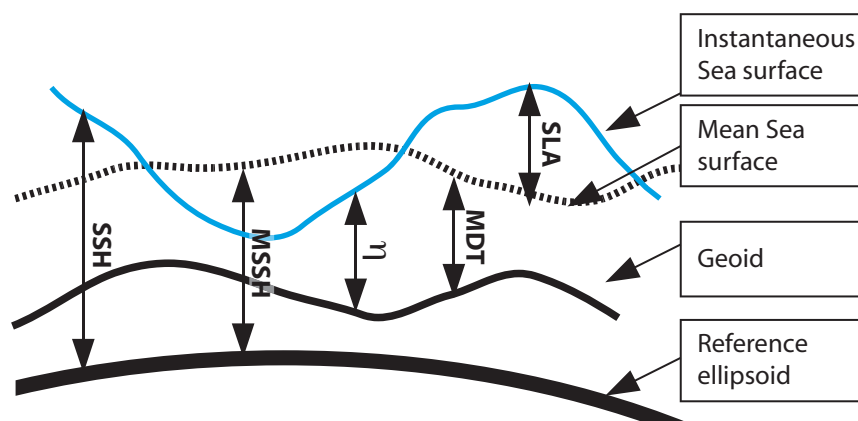


FIGURE 2.4: Decomposition of Sea Surface Height. (M)SSH: (Mean) Sea Surface Height. η : dynamic topography. MDT: Mean Dynamic Topography. SLA: Sea Level Anomaly

par CLS (Rio and Hernandez, 2004, Rio5)¹. Ce dernier utilise également des données marégraphiques et de flotteurs *in situ* pour son estimation. La précision de ces produits a encore augmenté avec le lancement de la mission GOCE. En théorie, l'existence d'une MDT basée sur les observations devrait permettre de corriger l'état moyen du modèle, en effet, si la MDT utilisée est issue du modèle, seule la variabilité des spatiale et temporelle incluse dans les anomalies altimétriques est assimilée. Cependant, comme on va le voir dans la suite, les choses ne sont pas si simples.

Plusieurs estimés de la MDT peuvent différer d'une constante arbitraire. Par exemple, sur la surface qu'il couvre (les très hautes latitudes ne sont pas présentes), la moyenne spatiale de Rio5 est 1,35m plus haute que la MDT correspondante du modèle, issu des réanalyses océaniques de l'ECMWF. L'ellipsoïde de référence étant arbitraire, il devrait suffire de rajouter un scalaire uniformément une bonne fois pour toute. Choisir cette correction telle que le moyenne spatiale des deux champs coïncide n'est pas forcément une bonne idée étant donnée qu'il y a de grandes différences à certains endroits qui peuvent être dues à des erreurs dans le modèle. Les inclure dans le calcul de la moyenne spatiale risquerait d'induire une dégradation de l'analyse dans certaines régions. Si ces régions sont d'importance pour l'application considérée, comme l'ouest du Pacifique pour la prévision saisonnière, alors l'utilisation d'une MDT basée sur les observation aura été néfaste. De plus, à regarder les structures spatiales de ces différences (figure 2.5) il est clair que la façon de déterminer cette correction n'est pas unique. En effet, la MDT Rio5 est plus basse dans le Pacifique et plus haute dans l'Atlantique que celle issue du modèle, ainsi tenter de la réduire dans un des océans l'augmenterait dans l'autre. Il n'y a pas de moyen simple d'évaluer l'équilibre optimal, et une erreur de quelques centimètres peut avoir un impact important sur la température et la salinité de la colonne d'eau. De plus en comparant Tap03 et Rio5 on s'aperçoit que l'incertitude sur ces produits est loin d'être négligeable.

Ceci illustre le fait qu'il est délicat d'utiliser une MDT basée sur les observations dans un système qui a développé librement sa propre MDT et qu'il est important d'utiliser une méthode plus sophistiquée, probablement incluant un schéma de correction de biais (Vidard (2001), Lea et al. (2008), Balmaseda et al. (2007b)). Une autre façon de procéder (qui n'est pas exclusive de la correction de biais) serait d'assimiler les différences de MDT durant la période antérieure à la présence des données altimétriques. On éviterait ainsi un saut lors de l'introduction de ces données.

Dans une configuration Atlantique-Nord à haute résolution, Birol et al. (2005) ont montré, en utilisant une MDT provenant de données GRACE, que cela permettrait d'améliorer la représentation des courants,

1. Il y a des produits plus récents, cf Maximenko et al. (2009) pour une revue de ceux ci

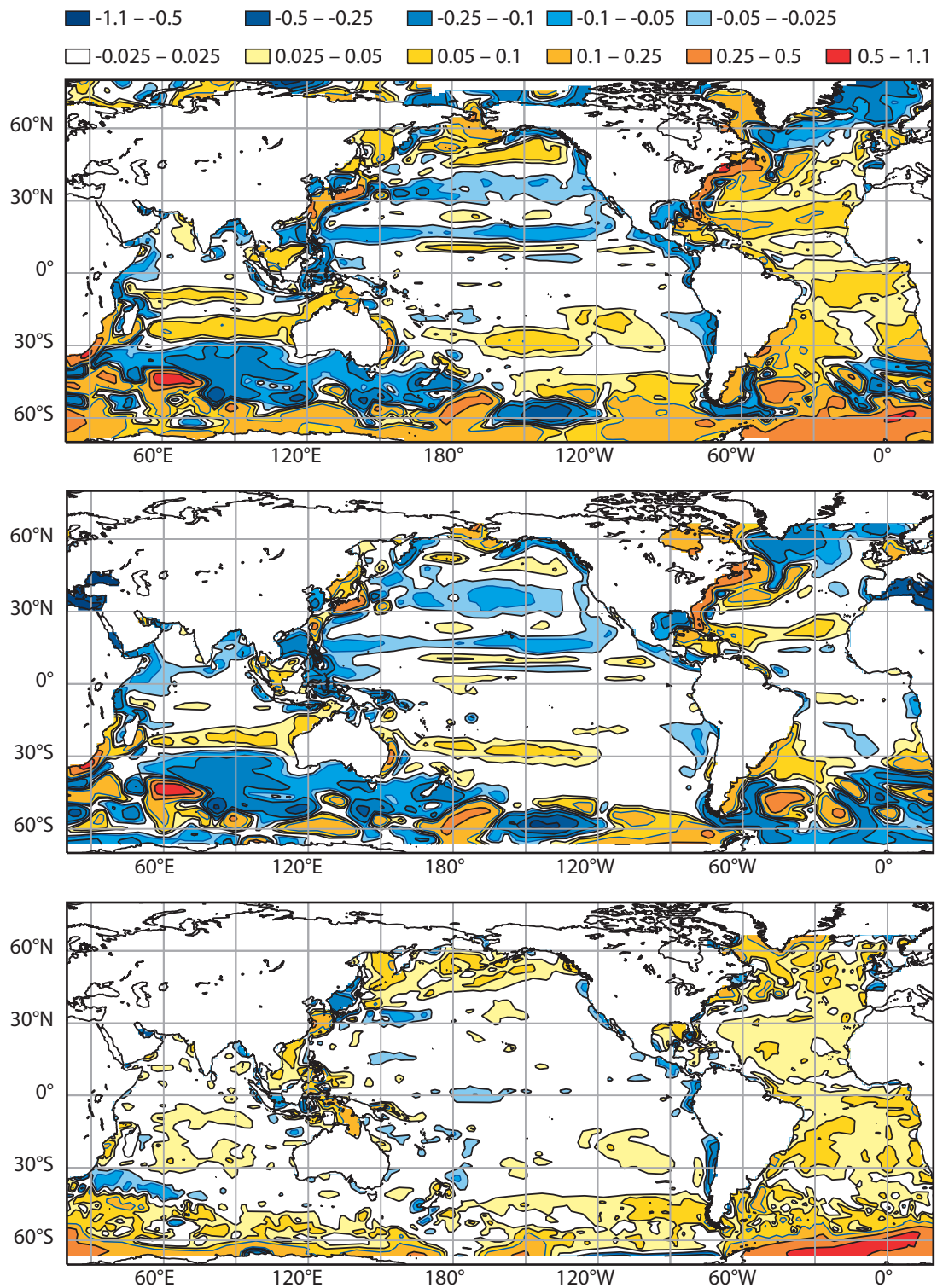


FIGURE 2.5: -Haut : Différence entre Rio5 et une MDT issue du modèle, -milieu : idem mais avec Tap03, - bas : différence entre Rio5 et tap03.

mais au prix d'une dégradation du champ de température. [Castruccio et al. \(2006\)](#), en utilisant un système à plus basse résolution, a semble-t-il trouvé un impact bénéfique en utilisant les données de GRACE. Cependant, bien que le modèle fut global, ils ont seulement utilisé la MDT provenant de GRACE dans la région du Pacifique tropical, suggérant par la même qu'ils rencontraient le même problème d'ajustement global.

Le choix fait dans [Vidard et al. \(2009a\)](#) et dans le system III de prévision saisonnière, fut de remplacer la MDT estimée par la MDT issue du modèle, ce choix du renoncement était sans doute judicieux à l'époque, compte tenue des contraintes de temps pour la mise en place d'un système opérationnel et des rigidités techniques de l'outil que nous utilisons, cependant il est évident que cela implique de jeter une partie importante de l'information. l'écart entre les différentes MDT s'est considérablement réduite et l'estimation des statistiques d'erreurs de celles-ci s'est affinée ([Maximenko et al 2009](#)), ce nonobstant, reste l'inadéquation de la MDT du modèle avant assimilation qui peut causer une correction forte dès que les données d'altimétrie commencent à être assimilées. ces erreurs de MDT restent de toute manière un biais dans le système, qui peut être pris en compte, en se basant sur des méthodes similaires à celles du contrôle d'erreur modèle systématique présentées section 1.3. Il s'agit en effet de considérer cette incertitude sur la MDT comme un biais des observations et de contrôler ce dernier. Cette approche avait été proposée comme extension des mes travaux de thèse sur l'erreur de modèle dans la dernière section de celle ci ([Vidard \(2001\)](#)), comme personne ne m'écoute, ce n'est que plus récemment qu'elle a été re-proposée et mise en pratique par [Lea et al. \(2008\)](#) et [Storto et al. \(2011\)](#), ou [Dee and Uppala \(2009\)](#) pour un problème similaire en météorologie. Elle consiste donc, pour une méthode de type variationnelle (mais une approche similaire peut être dérivée pour des méthodes de type filtrage) à contrôler un terme supplémentaire \mathbf{b} se situant dans l'espace de la MDT du modèle et à rajouter un terme dans la fonction coût pénalisant l'écart de ce nouveau terme de contrôle à une estimation \mathbf{b}^b *a priori*, issue par exemple d'une estimation précédente ou de la différence entre la MDT du modèle et une MDT observée. Ce nouveau terme peut évidemment être pondéré par des statistiques d'erreurs sur le MDT, obtenues à partir des différentes estimations disponibles de cette MDT.

$$J(\mathbf{x}_0, \mathbf{b}) = \frac{1}{2} \int_0^T \| \mathbf{y}(t) - \mathcal{H}[\mathbf{x} + \mathbf{b}] \|_{\mathcal{O}}^2 dt + \frac{1}{2} \| \mathbf{x}_0 - \mathbf{x}^b \|_{\mathcal{X}}^2 + \frac{1}{2} \| \mathbf{b} - \mathbf{b}^b \|_{\mathcal{B}}^2 \quad (2.3)$$

2.1.2 Tendence haussière du niveau de la mer

Une autre difficulté, plus due cette fois aux limitations du modèle et à la méconnaissance de certains forçages est causée par l'élévation observée du niveau global de la mer, comme représenté figure 2.6. Si localement, des changements durables de niveau de l'océan peuvent venir de phénomène divers comme des changements dans les vents dominants, (cf [Verlaan and Heemink \(2001\)](#) par exemple), au niveau de l'océan global il sont directement liés au changement climatique. Il y a deux facteurs expliquant cette élévation du niveau global des océans : l'expansion thermique (effet stérique) et la fonte des glaces continentale (apport d'eau douce).

Les changements "stériques" du niveau de l'eau sont causés par des changements de densité de la colonne d'eau impliquant une expansion ou une contraction de cette colonne. Ils sont principalement produits d'un réchauffement/refroidissement de la surface, et dans une moindre mesure des effets non linéaires de l'équation d'état. Les modèles non Bousinesq contiennent tous les processus océaniques qui peuvent agir sur le niveau de l'océan, en particulier ils contiennent cet effet stérique. Par contraste, les modèles faisant l'approximation de Boussinesq, comme c'est le cas de NEMO, conservent le volume d'eau, plutôt que la masse et ainsi ne peuvent représenter correctement ces expansions/contractions. L'effet stérique n'est ainsi pas représenté explicitement. Cette approximation n'induit pas une erreur significative dans la description de la dynamique du modèle ([Greatbatch, 1994](#)), a contrario, c'est un élément important quand on assimile les

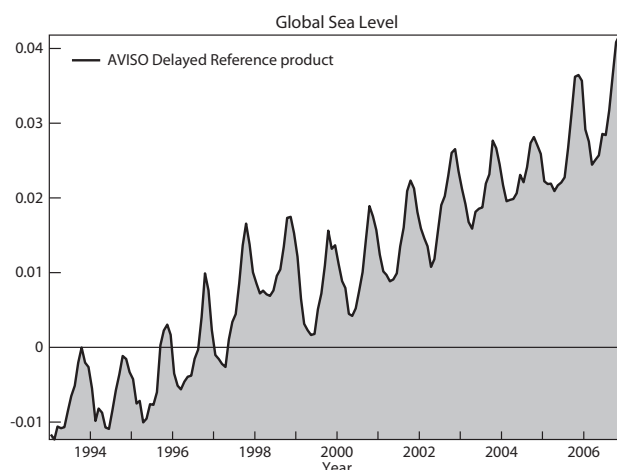


FIGURE 2.6: Série temporelle du niveau global de la mer mesuré par altimétrie (en mètres)

données altimétriques. En effet l'effet stérique est une composante importante dans le changement de niveau des océans. C'est encore plus vrai quand l'étude porte sur l'évaluation des effets du changement climatique sur le niveau des océans. Heureusement, la contribution stérique aux variations de hauteur d'eau peut être considérée comme uniforme spatialement et peut donc être diagnostiquée au travers du budget en masse de l'océan global. Dans [Vidard et al. \(2009a\)](#), la tendance haussière globale observée du niveau de l'océan est soustraite en enlevant le niveau global moyen des cartes d'observation d'anomalies avant d'assimiler ces dernières. Ensuite cette tendance est assimilée séparément comme une correction uniforme au budget de masse.

Les parts respectives dans la tendance haussière du niveau global de la mer dues à l'expansion thermique (effet stérique) et aux variations de masse (flux d'eau douce) sont encore débattues ([Church and White \(2006\)](#); [Balmaseda et al. \(2007d\)](#)). En principe les réanalyses océaniques devraient pouvoir aider à répondre à cette question. En effet, en combinant l'ébauche avec les informations de sous surface, on devrait être en mesure de calculer l'effet stérique. D'un autre côté, en comparant la tendance dans le niveau global de la mer observée avec la contribution stérique estimée ci-avant, il sera possible de donner une estimation de la composante due à l'apport de masse.

2.1.3 Une nouvelle génération

L'exemple de l'altimétrie illustre que même dans un cas relativement favorable (on assimile une quantité faisant partie des variables d'état du modèle), la mise en pratique peut s'avérer délicate. Les contraintes opérationnelles : ressources limitées (surtout en océanographie) et impératifs temporels ne permettent pas de toujours mettre en place des solutions optimales, même quand elles existent. C'est un aspect important dans la relation académiques-opérationnels.

L'histoire de l'assimilation de l'altimétrie ne s'arrête pas là, même si le financement et la pérennisation de tels instruments sont toujours incertains sur le long terme, de nouveaux satellites vont prochainement être lancés, soit en remplacement soit en addition de ceux en mission actuellement (Fig 2.7). Les données vont devenir plus nombreuses et de meilleure qualité, en tirer le meilleur parti reste un défi. En particulier, un nouveau type d'instrument à large pinceau (120km), à traces resserrées (+/-10 km contre 100 actuellement) et à haute résolution (20-60m) est prévu sur le satellite Surface Water and Ocean Topography (SWOT), ce genre de données requerra certainement un traitement différent, peut-être s'approchant plus de l'assimilation

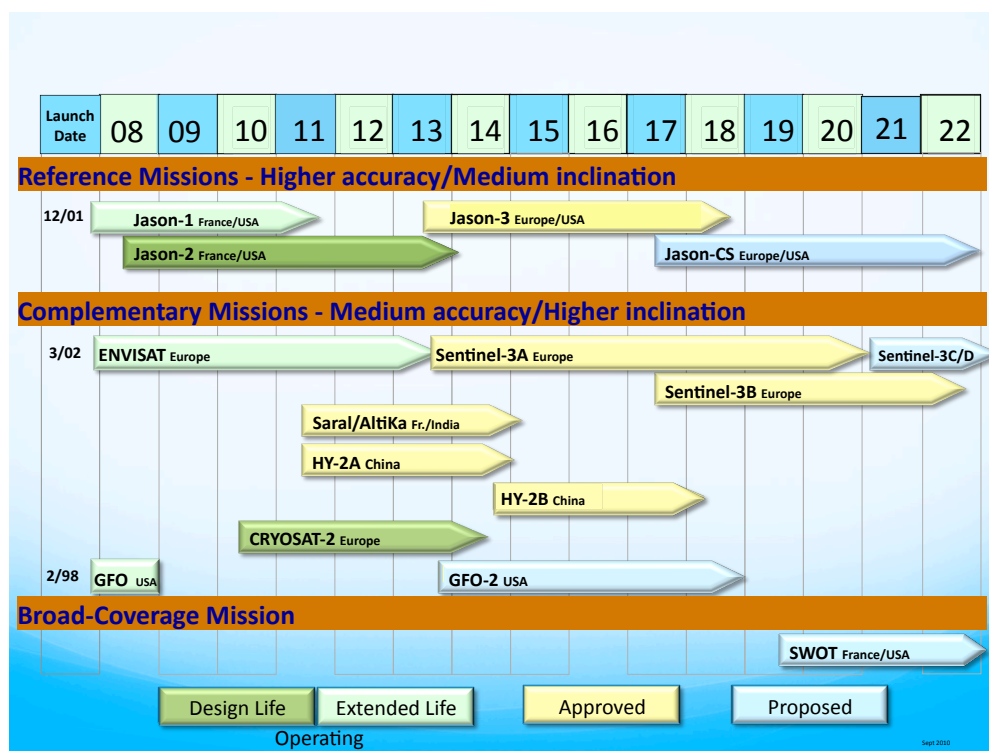


FIGURE 2.7

d'images (cf section suivante). En ce sens il ouvrira la voie à une nouvelle révolution (sans jeu de mot) pour l'océanographie spatiale. Il ouvrira également d'intéressantes possibilités pour les systèmes de prévision d'hydrologie terrestre et donc de nouveaux horizons pour l'assimilation de données.

2.2 Assimilation d'images

Cette section est issue des notes du cours que j'ai donné à l'école d'été 2012 des Houches², il a été préparé en collaboration avec O. Titaud et se base sur nos travaux communs ainsi que des travaux de thèse d'Innocent Souopgui et Vincent Chabot que j'ai eu le plaisir de co-encadrer (et ai toujours pour ce dernier) avec F.-X. Le Dimet et M. Nodet respectivement, et ce dans le cadre des projets ADDISA et GeoFluids. Ces notes ont été rédigées en anglais, d'où un changement temporaire de langue d'expression dans ce manuscrit.

2.2.1 Motivations

Since the end of the seventies, many satellites have been launched to improve our knowledge of the atmosphere and of the oceans by observing the Earth. The geostationary ones provide photographic images of the earth system among other data. They provide sequences of such images that show the dynamical evolution of certain meteorological or oceanic "objects": fronts, clouds, eddies, vortices, etc. Dynamics that the human vision can easily detect in this kind of image sequence clearly has a strong predictive potential

2. <http://houches2012.gforge.inria.fr/>

(Fig 2.8). This aspect is clearly favoured by the fact that these data, contrary to many other measurements, are dense in space with a resolution close to one kilometre and in time with a full image every 15mn for current METEOSAT satellites and every 10mn (and even every 2.5mn for Europe only) for the upcoming third generation. It implies a huge amount of data which can be seen as an asset but also induce difficulties for the assimilation system for coping with such amount of data. In practice only a tiny percentage ($\approx 3 - 5\%$) of total satellite (from polar orbiting and geostationary) data are used in operational NWP systems and images are only used through pseudo observation (see later) and given low confidence within the assimilation systems. Considering the cost of the satellite observing systems (the cost of the launch of the Meteosat Third Generation is estimated at around 2.5 billion Euros) and of the infrastructures required for the collect of the data itself, improving their impact on forecasting systems is an important topic of research.

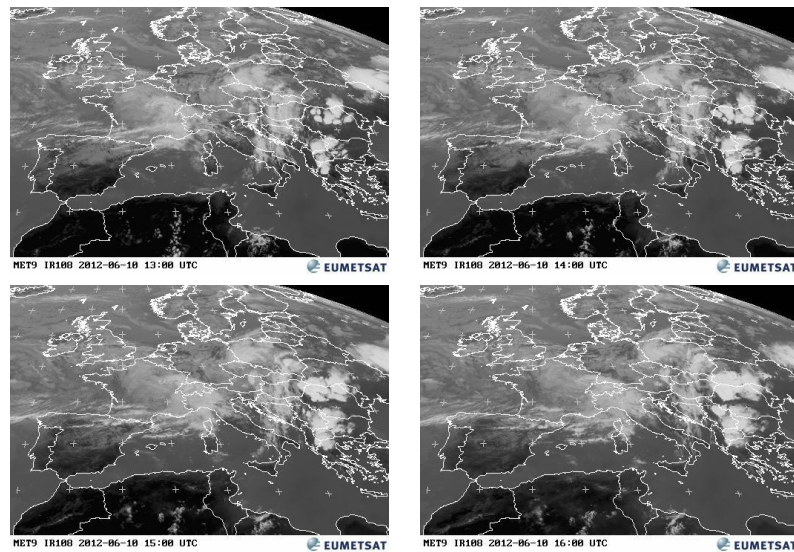


FIGURE 2.8: sequence of Meteosat images over Europe the 12th of June 2010 between 13:00 and 16:00 UTC (source Eumetsat)

The purpose of this section is therefore to present a brief introduction to image sequences assimilation in a geophysical context. As said before, this is still a topic of research and only basic options are used in operational systems. In relation, two classes of image assimilation techniques can be considered:

- Pseudo-Observation: an apparent velocity field is estimated from an image sequence using certain image processing techniques. This estimated field can then be used as indirect observations in a classical assimilation scheme. This is what is currently done in a meteorological context (Schmetz et al., 1993).
- Direct Image Assimilation (DIA): images or image sequences are assimilated through an appropriate observation operator directly into the Data Assimilation System. This idea follows a general trend to avoid pre-processing of observations before its assimilation. This is not yet used operationally and a selection of possible metrics is presented.

Using images or image sequences together with numerical models may present several difficulties: First the state variables of the numerical models (e.g. wind, temperature, pressure, humidity in atmosphere modelling and current velocities, temperature, surface elevation, salinity in ocean modelling) are not directly measured by satellites. What is observed is more or less linked to those variables by diagnostic equations that will have to be included in the assimilation procedure. Moreover, the physical processes that are obser-

ved are not always taken into account in the model (e.g. local convection in atmospheric modelling, ocean colour in ocean modelling). These two aspects are not specific to images and may happen for other kinds of observation as well but they are likely to be systematic in the case of image sequences.

More specifically, images are bidimensional informations whereas physical processes of geophysical fluids are three-dimensional. From this point of view, a major difference between the ocean and the atmosphere comes from their radiative properties. In the first case images give informations about the ocean surface whereas in the second case images integrate all the radiative profile of the observed atmosphere column. Besides, certain meteorological structures (e.g. some kind of clouds) are located in specific layers. Assimilating informations about the dynamical evolution of these structures needs additional informations about the altitude of the observed processes (using temperature profile measurements for instance). When the radiative properties of the ocean make the images give information about the ocean surface only (typically about a few centimetres, up to a few meters in some cases), the corresponding observed processes may not be well depicted by the model. Additionally satellite Images can be of relatively poor quality, this is mostly true for ocean surface images very often partially occulted by clouds. Moreover they are most of the time composite images: it takes Meteosat 15 mn to take a whole image chunk after chunk and the surface SST images come from polar orbiting satellites and could take a few days for full covering a given area (Fig 2.9). In a sequence, the luminosity of individual images in a sequence may vary according to the time of the day and only apparent motion is observed that could lead to aliasing problems. Another specificity of images is that, as for human vision, we can presume that the pertinent informations coming from an image are mainly brought by its discontinuities or high gradients. Unfortunately, numerical models have a tendency to smooth these discontinuities out.

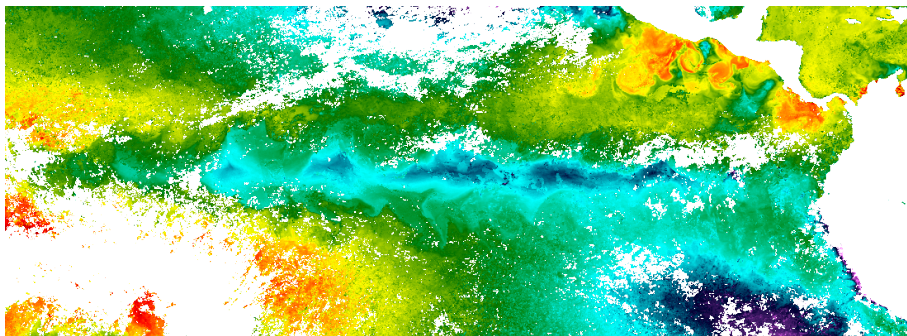


FIGURE 2.9: Infra red composite image from Metop 21st of January 2008 (source Ifremer)

Finally, as aforementioned handling the massive amount of data included in an image sequence is particularly challenging

In order to tackle the above-mentioned difficulties, one has to carefully design the assimilation system. It should be able despite these difficult conditions to extract from the observations the information about the dynamics of the system. The difficulty in building such system lies in the adequate definition of the observation space and the related metrics and consequently through the construction of the observation operator.

2.2.2 Images: Level(s) of Interpretation

Before really considering the assimilation of images, let us better define what an image is in this context. For the human eyes it is structured information, dense on the domain it covers; for a computer it is an array of pixels and from a mathematical point of view... it can be many things.

But basically an image is considered as a function of two variables

$$\begin{aligned} I &: \Omega \subset \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R} \\ (x, y) &\mapsto f(x, y) \end{aligned} \tag{2.4}$$

whose definition can be of very different nature (see below)

Numerical raw images

The acquisition device converts the radiations coming from the observed scene into an electric signal that is digitalised to form the *raw image*. A numerical raw image is an array of pixels whose values represent the quantity of photons the corresponding photo-site received during the acquisition procedure. Raw images pixels can then be considered as radiance measurements. However, from a *mathematical* point of view, an image is commonly considered as a real valued function of two continuous real variables. This allows the use of powerful mathematical tools developed by the image processing scientific community. A *numerical image* is the discrete version of a mathematical one that is finally manipulated by computers.

Levels of interpretation

An image can be considered at different level of interpretation:

At the *pixel level*, an image of a fluid in motion is a digital instantaneous expression of the state of the fluid through the radiation that it reflects (visible light) or emits (infrared light). It is then a snapshot of an *apparent state*, which may involve other physical quantities than what the model simulates, like tracers (e.g. Cloud cover and aerosols in the atmosphere, chlorophyll in the ocean). The radiances are generally not part of the studied system and the dependence between this radiation and the state variables is not often trivial. For instance, *cloud cover* is an intricate function of the state variable of meteorological models: clouds reflectance and intrinsic infrared emission depend both on thermodynamics (temperature and humidity) and on the micro-physical processes that occur inside the clouds (involving water, ice, snow and the size of the particles). At this level, information contained in images is not structured and represents a huge amount of data: typical satellite images are about 5000×5000 size, that is 25 millions of pixels. This number should be multiplied by the number of frames in the observed sequence.

At the *analysis level(s)*, an image is symbolically described by its contents, typically by the objects and their characterisation (geometry, layout, etc.) that appear in the observed scene or on its mathematical description (see further). This kind of interpretation is used for high-level vision problematics as motion estimation, edge detection, pattern recognition, etc. The advantage of this level is that it may less depend on the effective radiance emission and it allows emphasising on the structures the image contains (and even their evolution in an image sequence) rather than the actual value of individual pixels.

A good illustration of these different levels of interpretation is the two images from two different MODIS sensor shown in figure 2.10. At pixel level they represent two very different and somewhat independent quantities of very different nature (Temperature and Chlorophyll) while at a higher level of interpretation, both images give a similar kind of information about the dynamic of the system. Indeed, one can easily notice the vortices and fronts that are present in the images.

Mathematical definition

There exist several mathematical definitions for images and it is not the aim of this section to describe them all and only a small subset will be discussed in the following. They essentially depend on the way the image needs to be processed. We can cite the multi-scale decomposition approach which comes from the

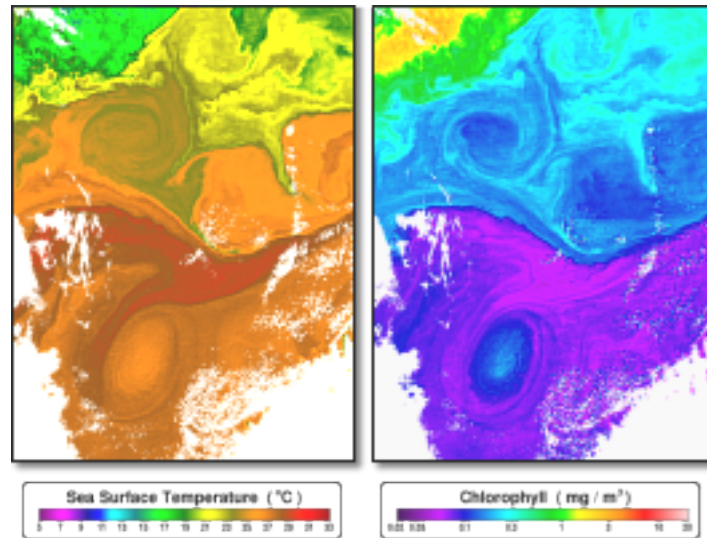


FIGURE 2.10: Images over the Gulf Stream the 4th of August 2005 from Aqua satellite (MODIS radiometer). Ocean Sea Surface Temperature (IR channel, left) and chlorophyll concentration (visible channel, right). (Source : NASA)

signal processing theory (*wavelet* (Mallat, 1998) and the more recent *curvelet* (Candès and Donoho, 2003) transforms). Another approach uses the Partial Differential Equation framework (Aubert and KornProbst, 2006) where the image is considered as a function of two variables verifying a PDE equation that depends on the considered image-processing problem. Certain applications use a stochastic approach as well to define images as a combination of an "ideal" image (most often deterministic) and a random function (Geman and Geman, 1984). For a good review of image analysis and the underlying mathematical definition, see (Chan and Shen, 2005).

In the DA framework, the level of interpretation together with the mathematical definition of the image may be represented by the image-type observation space and its corresponding observation operator (see Section 2.2.4).

2.2.3 Current Use of Images in DA: Pseudo Observation

Sequences of images are widely available in meteorology, oceanography, hydrology, astrophysics, and even medicine. Historically, they have been mainly used for a "by eye" analysis from experts in these fields. Yet, since the mid 80's, research has been carried out to derive velocity fields from the sequences, with applications for fluid dynamics mainly (and very recently for movie compression and medical applications). One way to assimilate images is to use the velocity field derived from some image processing techniques as pseudo-observation of the studied system. This class of Motion Estimator, directly based on image processing techniques, aims at estimating the velocity field that would "transport" one frame to another or one part of the frame to the other frame.

There are several ways to extract a velocity field from a sequence of images. They are mainly working on two successive frames, even though some extensions allow for the use of more frames at the same time. Hereafter is a short list of the most common methods.

Bogus

The simplest method to assimilate images is to detect specific features in the images and add or displace them in the system. Historically both steps of this process were done manually by forecasters and modellers, but recently some effort has been carried out to do the detection automatically, like for instance with the dry intrusions that are important precursor of extreme event (see figure 2.11 and Michel and Bouttier (2006)) as well as for the assimilation itself *e.g.* with the use of pseudo observation of potential vorticity structures Michel (2011).

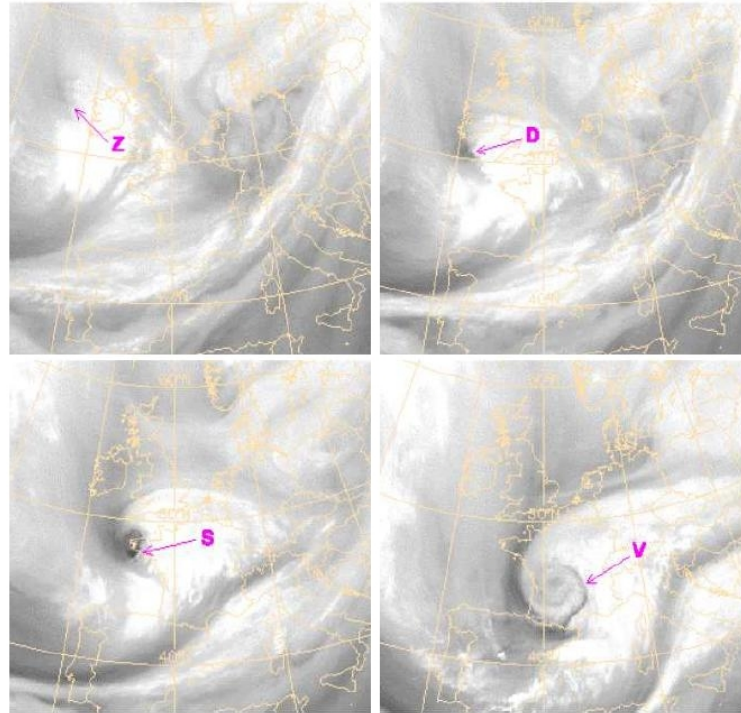


FIGURE 2.11: Evolution of a dry intrusion (source Météo-France)

Statistical methods

They were set up for Particle Image Velocities (PIV) and Particle Tracking Velocities (PTV) experiments *i.e.* for lab experiments in fluid mechanics where particles are added to the fluid in order to follow the flow (Adrian, 1991). In this kind of method two successive images are divided into small areas and cross-correlation techniques can be used to match each individual areas of the first image with one of the second image.

$$C = \frac{\sum_{i,j}(T_1(i,j) - \bar{T}_1)(T_2(i,j) - \bar{T}_2)}{\sigma_{T_1}\sigma_{T_2}} \quad (2.5)$$

with T_1 and T_2 represent the spatial distribution of grey values in the first and second image templates, respectively. The T_2 corresponding to the highest C is elected (if C is larger that a given threshold, typically 0.8).

It is then possible to create a velocity field that would transport all the areas of the first image to the second one. They are easy to implement, but they can be quite expensive when the number of areas increases.

Since the resulting velocity vectors are based on cross-correlating the intensity distributions over small areas of the flow, the resulting velocity field is a spatially averaged representation of the actual velocity field. This obviously has consequences for the accuracy of spatial derivatives of the velocity field, vorticity, and spatial correlation functions that are often derived from PIV velocity fields. PIV methods will obviously not be able to measure vertical velocities. These components might not only be missed, they can also induce errors in the estimated horizontal velocity fields. There also exist Stereoscopic PIV, which uses two cameras to estimate all three velocity components (if the observed media allows for it) but they are for obvious practical reason only available for laboratory experiment for now.

This is the kind of method that is actually currently used in NWP to derive winds from the satellite images from geostationary satellite (METEOSAT, GOES and GMS), with the so-called AMV (Atmospheric Motion Vectors). The clouds are used in this process instead of the particles in classical PIV methods and due to the peculiar nature of these tracers the images require two pre-processing steps prior to the wind estimations. First, since the tracers (clouds) are not uniformly spread across the image one has to detect, in each area, features of interest that would be tracked from one image to another with methods generally based on the gradient of the images. A lot of quality control procedures are then applied to the selected features to retain only those that will be traceable. For instance, the gradients have to be large enough, cover a significant portion of the area, avoid areas more than two different gradient amplitude, ...

Second, images are 2D representations of a 3D system and therefore one has to allocate a height at a given tracked feature (with the reasonable assumption that clouds will not move up and down between two images). Historically (and still used as a fallback) the window channel brightness temperatures within the target area were analysed, and a mean value for the coldest 20% of the sample was used to represent the temperature at cloud top. This temperature was then compared to a numerical forecast of the vertical temperature profile to arrive at the height of the cloud. More recently AMV producers use the so called CO_2 Slicing where the ratio of the IR and CO_2 absorption can be linked to the pressure at the location of the observed feature or the H_2O intercept that uses the fact that the radiances from a single-level cloud deck for two spectral bands vary linearly with cloud amount. Both CO_2 slicing and H_2O intercept require a model forecast for calibration. This step is an important source of errors in the derivation of AMVs, moreover these errors are correlated with the model error since a model forecast is used in the processes. This last aspect is an important breach of the common assumption of uncorrelated errors between model and observation that is made by all the main data assimilation algorithms.

Once this is done PIV methods can be used to derive wind fields on several altitudes. In order to avoid too much noise in the resulting wind field three images are used instead of two to derive 2 wind fields (from image 1 to 2 and from a image 2 to 3) and only the sum of the resulting vectors that do not disagree are kept for the final product. (a complete description of these processes can be found in [Schmetz et al. \(1993\)](#) and [Nieman et al. \(1997\)](#)). Figure 2.12 represents the operational AMVs derived by Eumetsat using the infrared image sequence showed in figure 2.8.

Due to their really indirect nature and the complexity of the pre and post processing, describing the errors associated to such wind data is not straight forward, in particular they are correlated so complex observation error covariance matrices have to be built or the errors have to be significantly inflated, and therefore it will reduce their impact. [Bormann et al. \(2003\)](#) found statistically significant spatial error correlations on scales up to about 800 km, and moreover they are strongly anisotropic.

Optical flow

Variational methods are probably more adapted to the treatment of the large image sequences available in geophysics even though they are seldom used in practice. The most common one is usually called *opti-*

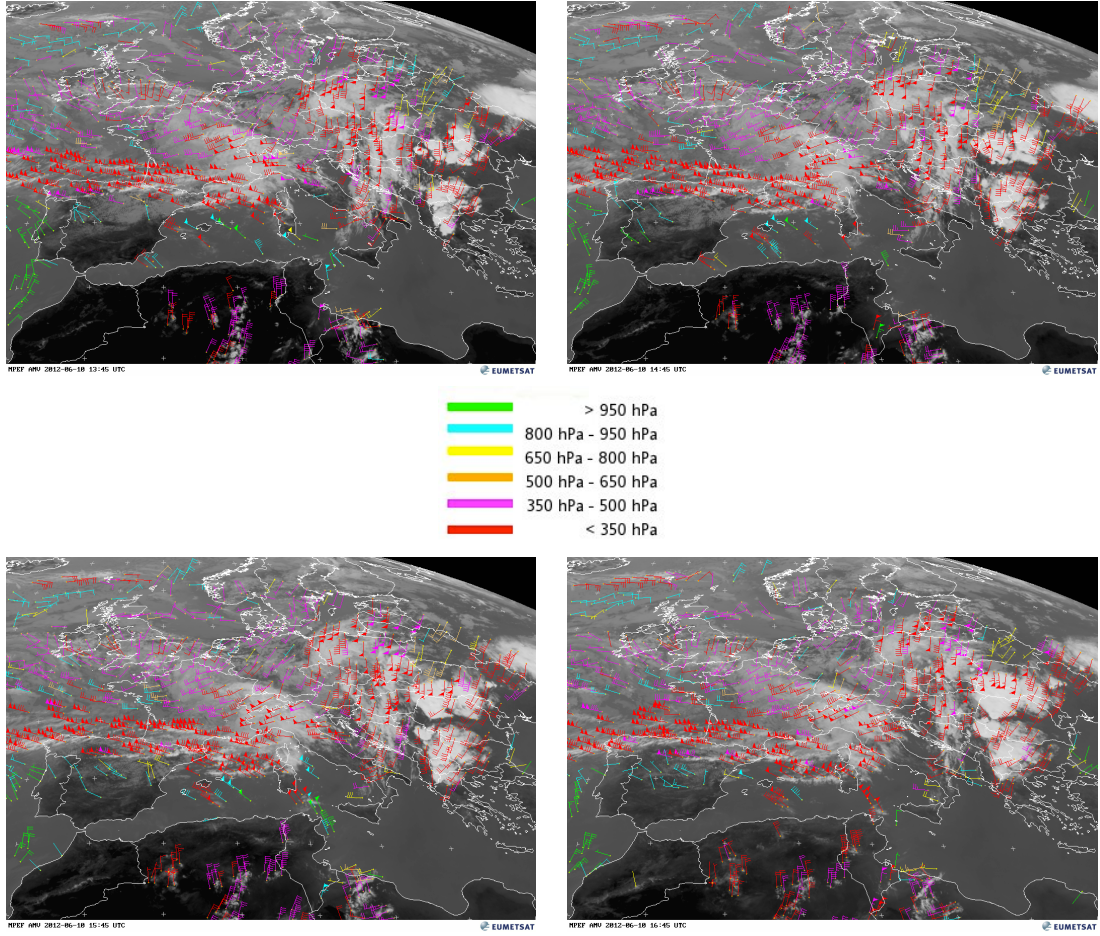


FIGURE 2.12: Sequence of Meteosat's derived Atmospheric Motion Vectors over Europe the 12th of June 2010 between 13:45 and 16:45 UTC (source Eumetsat)

cal flow. This classical approach in computer vision is based on the conservation of the global luminance between two images (Horn and Schunck, 1981) and they aim at finding the velocity field $\mathbf{w} = (u, v)$ that transport two successive images I_1 and I_2 one to another.

$$I_1(x + u(x, y), y + v(x, y)) = I_0(x, y) \quad (2.6)$$

A vector field satisfying this equation is not unique (aperture problem) and a strict conservation is not realistic in general. So it is transformed into a minimisation problem:

$$J(\mathbf{v}) = \frac{1}{2} \int \|I_1(x + u(x, y), y + v(x, y)) - I_0(x, y)\|^2 dx dy + \frac{1}{2} \alpha R(u, v) \quad (2.7)$$

The choice of the regularisation term depends on the application field (Vigan et al. (2000) for oceanography, Amodei and Benbourhim (1991) for wind field). A good review of possible R is available in Auroux and Fehrenbach (2010) and some thoughts about preconditioning in Souopgui (2010). The advantage of such method over PIV is that it can provide a denser velocity field for a fraction of the computing cost (see Figure 2.13) however the adjustment or the relative weight of the two terms of the cost function (*i.e.* the value of the parameter α) and the choice of the regularisation term are far from trivial (Souopgui, 2010).

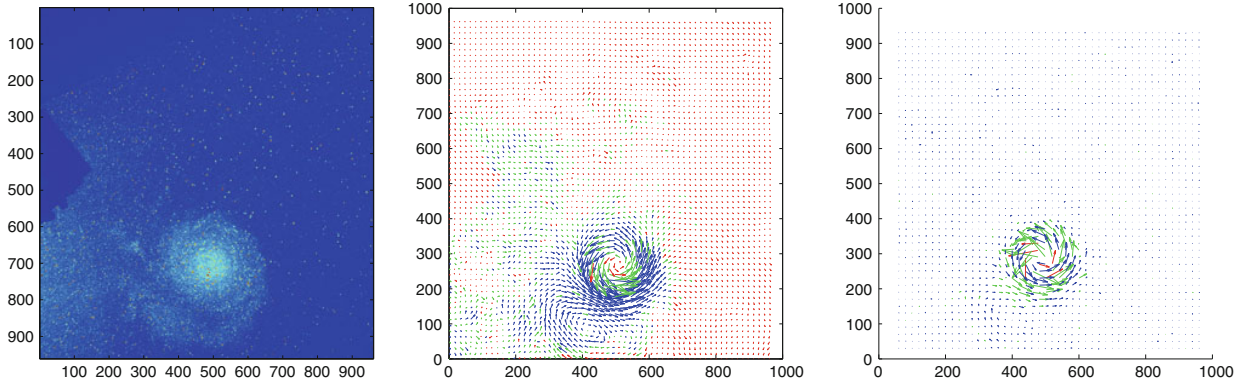


FIGURE 2.13: From left to right: First original image, velocity field derived from optical flow, velocity field derived from PIV. Blue, green and red: high, medium and low confidence respectively (from [Auroux and Fehrenbach \(2010\)](#))

Once again this motion estimator is meant to be used for a pair of images, however as for statistical (but in a more natural manner) it can be extended to use more frames altogether in order to improve consistency, by adding a governing equation for \mathbf{v} to the optimality system generating the pseudo observations.

As above, it assumes the conservation of light and the image pixels are transported from one image to the other by the velocity field to be estimated. In order to improve the temporal consistency of the estimated velocities, this method requires a simple temporal evolution model \mathcal{N} . Along with the transport equation of the pixels (2.7) it forms:

$$\begin{cases} \frac{\partial I}{\partial t} + \nabla I \cdot \mathbf{v} = 0, \\ \frac{\partial \mathbf{v}}{\partial t} = \mathcal{N}(\mathbf{v}), \quad I(0) = I_0, \quad \mathbf{v}(0) = \mathbf{v}_0. \end{cases} \quad (2.8)$$

The pixel values of the images I_{obs} are then assimilated in the previous model using variational data assimilation (see Section 2.1) where the control variable is the initial condition \mathbf{v}_0 of \mathcal{N} .

$$J(\mathbf{v}) = \frac{1}{2} \sum_{i=1}^{m-1} \int \|I_i(x + u(x, y), y + v(x, y)) - I_{i-1}(x, y)\|^2 dx dy + \frac{1}{2} \alpha R(u, v) \quad (2.9)$$

The advantage of this approach is that it takes into account all the images of the sequence at once and therefore can cope with missing data on some of the frames. This has been applied, using a constant velocity ($\mathcal{N} = 0$) in ([Korotaev et al., 2008](#)) to derive ocean surface velocity fields from SST on the Black Sea.

Advantages and limitations of the pseudo-observation technique

The main advantage of such approach is simplicity; indeed, even though the algorithm use to produce pseudo velocity observations may be sophisticated, the use of the pseudo observation in an already existing DA scheme is relatively straightforward. There is no issue of differentiation and no heavy development is generally needed since this processing is done off-line.

However, despite their relative simplicity for implementation and their rapidity, pseudo-observation techniques based on motion estimation may suffer from some limitations. First, a frame-to-frame motion

estimator does not take into account a sequence of more than two images. Due to the lack of consistency in time, it cannot capture the dynamical evolution of the image during the entire assimilation window. One of the consequences is that it cannot deal with missing data, which can be quite frequent in some of the applications targeted here. Missing data can be due to a glitch in the observing system or an obstacle (e.g. clouds). Solving 2.9 may mitigate this disadvantage, yet the model \mathcal{N} of the temporal evolution of the velocity only depends on \mathbf{v} and therefore is likely to be a somewhat poor representation of the real behaviour of the velocity.

Neither of these two techniques takes into account physical information about the underlying physical processes observed in the images. Furthermore some assumptions like the conservation of the luminance are clearly not valid for fluid flow images showing broken up structures. Finally, velocity fields obtained through the pseudo-observation approach are apparent velocities that can be quite different from the actual velocities. For instance, the absence of a tracer in a subset of an image sequence will lead to null apparent velocities in that area. Such techniques then need to provide error maps to account for this kind of error in the interpretation of the observation. Therefore, the definition of the observation operators and the specification of the corresponding observation error matrices are likely to be a difficult task [Bormann et al. \(2003\)](#).

For all these reason an important research effort is being carried out on that topic in order to define a methodology direct image data assimilation that would combine the information coming from image sequences and the ones coming from the numerical model, a-priori knowledge and classical observations in a more consistent way (e.g. without the need of pseudo observations).

2.2.4 Direct Assimilation of Images

If one wants to avoid going through the pseudo observation route the main difficulty is to define a way of measuring the discrepancy between the model outputs and the images. That is to define the observation space \mathcal{I} and the associated distance $d_{\mathcal{I}} : \mathcal{I} \times \mathcal{I} \rightarrow \mathbb{R}$, and the observation operator $H : \mathcal{X} \rightarrow \mathcal{I}$. There are many ways to define this triplet from very basic to highly sophisticated. For instance, to define \mathcal{I} we could consider

1. Frequency characteristics: in that case, \mathcal{I} is the image space of a multi-scale transformation such as wavelet and curvelet transforms. This approach is interesting for its well-known efficiency in data compression, denoising and edge extraction.
2. Geometric characteristics: certain image analysis techniques can detect certain geometrical features in the image (e.g. by means of active contours). This feature may take the form of a parametrised curve in \mathbb{R}^2 which could be assimilated as Lagrangian observation. This approach will necessitate high-level image analysis.
3. Qualitative characteristics: in geophysics, there are several recognisable structures connected to physical phenomena (e.g. cyclone, front, etc.). We could define \mathcal{I} as a dictionary of such couples observable structure / physical phenomenon. More specifically in meteorology, clouds may give information about the flow or the physical processes where they are located. For instance, the particular shape of the stratus clouds is due to a stratified flow whereas cumulo-nimbuses are known to involve strong convective processes.

In order to choose the above-defined triplet, one may start from the observation operator. $H_{\mathcal{I}}$ has to extracts structures from model state variables (possibly in a differentiable manner) and this can be done, either by creating a synthetic image from the model output and then perform the same extraction on both observed image and synthetic image in order to define the observation space and its associated distance, or by directly extracting these structure from the model output.

In the following two examples of such definition are presented, one using the synthetic observations route, the second directly extracting dynamical features from the model output and matching them to the one from the images.

Synthetic images and multi-scale transform

Features in geophysical images such as eddies or filaments may correspond to tracers in the observed system. A tracer can be defined as any quantity that is transported by the fluid flow. For instance, potential vorticity, temperature, salinity can be considered as atmospheric and ocean tracers. Modelling and simulating their evolution from the model state variables may help to define the Structures Observation operator. Indeed, we may use their description maps (e.g. concentration maps) as a *Synthetic Image Sequence* from which comparison with images (in a sense that remain to be defined) can be performed.

In other words, if \mathbf{q} is the tracer concentration, one can model its evolution by transporting it through the model velocity field using a standard transport equation:

$$\frac{\partial \mathbf{q}}{\partial t} + \nabla \mathbf{q} \cdot \mathbf{u} = 0 \quad (2.10)$$

where \mathbf{u} is the model velocity. This allow to produce a field that will be comparable to observation from the model output; once again, as for optical flow, this equation can be different, depending on the observed quantity. This way of proceeding requires the availability of the initial tracer field. It can be done thanks to an image at the beginning of the time period, but it cannot cope with missing data. Alternatively, one can augment the state vector with the concentration \mathbf{q} and control its initial condition (or filter its trajectory depending on the assimilation scheme) along with the state vector's.

Yet certain features of interest in geophysical fluid images do not correspond to any tracers (such as waves or dry intrusions in atmosphere images). In that case, other techniques should be developed to extract structures from model outputs without synthetic image sequence production.

From this observation operator there are plenty of possible choices for the observation space. The simplest of them being L^2 with the classical norm: $d^2(\mathbf{x}, I) = \|H(\mathbf{x} - I)\|_{L^2}^2$. In the discrete case, it means comparing the synthetic image ($H(\mathbf{x})$) to the observed one, pixel by pixel independently. [Papadakis and Memin \(2008\)](#) showed that assimilating pixel levels directly gives a better result than using an optical flow-based pseudo-observation approach. Also in [Corpetti et al. \(2009\)](#) such formalism is used to assimilate images of pressure differences into a three-layer simplified atmospheric model. Nevertheless, we could point out two possible limitations of the use of pixel level for realistic applications in geophysics: first for large images, the number of observation will become enormous (one per pixel) making difficult any assimilation, second it cannot differentiate errors in placement with error in amplitude of the observed/simulated features.

This was the motivation for the approach proposed in [Souopgui \(2010\)](#); [Titau et al. \(2010\)](#) and based on a multi-scale decomposition of both synthetic and observed image. In the present case a curvelet transformed was used, but similar results can be obtained using other multi scale decomposition like wavelet for instance. It actually uses the fact that any image can be decomposed as:

$$I = \mathcal{C}(I) \equiv \sum_{j,k,l} \langle \varphi_{j,k,l}, I \rangle \varphi_{j,k,l}, \quad (2.11)$$

where $(\varphi_{j,k,l})_{j,k,l}$ are the elements of the curvelet frame. The inner product $\langle \cdot, \cdot \rangle$ is the usual $L^2(\mathbb{R}^2)$ one. The curvelet transform is a linear isometry and then its adjoint is given by the inverse transformation (reconstruction). This is a convenient feature for practical use in variational data assimilation framework. A threshold \mathcal{T} is then applied to this decomposition (only a subset of (i, j, k) terms are kept) in order to extract

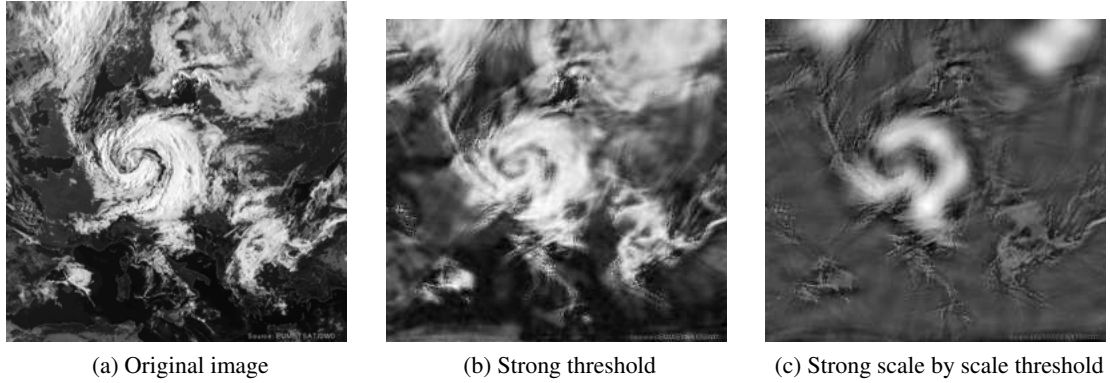


FIGURE 2.14: Original image (2.14a) and its curvelet / threshold / reconstruction product for a strong threshold applied to all the coefficients (2.14b) or independently scale by scale (2.14c) both retaining 10% of the coefficients in the curvelet frame

the dominant features of the images. Figure 2.14 shows the effect of two different thresholding techniques on the reconstruction of a cloud image. The observation becomes:

$$\mathbf{y} = \mathcal{T} \circ \mathcal{C}(I) \quad (2.12)$$

In order to compare with the model, the same transformation is applied to the synthetic image obtained from the model output

$$\mathcal{H}(\mathbf{x}) = \mathcal{T} \circ \mathcal{C}(\mathbf{q}) \quad (2.13)$$

with \mathbf{q} linked to \mathbf{x} by equation 2.10. Then we can define a semi norm on \mathcal{I} from the term-by-term product of the curvelet coefficients that survived the threshold function.

The choice of the threshold \mathcal{T} is of crucial importance as shown in Souopgui (2010) and illustrated by the figure 2.15, which compare the velocity field obtained by direct image assimilation in a shallow water model of images coming from a laboratory experiment performed at the CORIOLIS turntable (as a watermark behind the vector field). This example emphasises on the robustness of the multi scale approach to noise on the data. The L^2 norm (2.15a) gives a noisy result with numerous small displacements, which are actually artefacts of the bad quality of the observation images. On the contrary a hard threshold (2.15b, where all the curvelet coefficient below a given threshold are discarded) will have the tendency to underestimate the velocity field, while using a finer threshold (2.15c where a strong threshold is applied scale by scale independently), retaining the same number of coefficients will lead to a better reconstruction of the isolated vortex, with the right amplitude and without the unwanted small currents. It has to be noted that the convergence of the minimisation using the L^2 comparison tends to be much slower than that of the curvelet based comparison. This is due to the larger amount of observation (number of pixels vs number of curvelet coefficients) and to the stronger correlation between the pixels of a same image. This leads to the degradation of the conditioning of the minimisation problem (Haben et al., 2011). For longer assimilation windows, the L^2 norm based approach does not even converge while the truncated curvelet based one still manage to find a minimum. These results are obtained using images of a relatively modest size (256×256), this problem is likely to be amplified when switching to Meteosat-type images (5000×5000).

Moreover, the use of the pixel level may then lead to some increase in the representativity error. Note that small-scale structures may be the consequence of meso-scale dynamical processes. This is particularly true

for mixing tracer processes where spirals and filaments may come from the variability of meso-scale velocity field as it is clearly shown in [Lehahn et al. \(2007\)](#). In a pixel basis, those filaments are not represented as coherent structures while another mathematical modelling could avoid this problem.

Compared to the classical L^2 norm, the use of higher level of interpretation reduces significantly the size of the observation vector (typically only a few percent of the coefficients are kept). Additionally it tends to filter out small scales that cannot be represented by the model anyway.

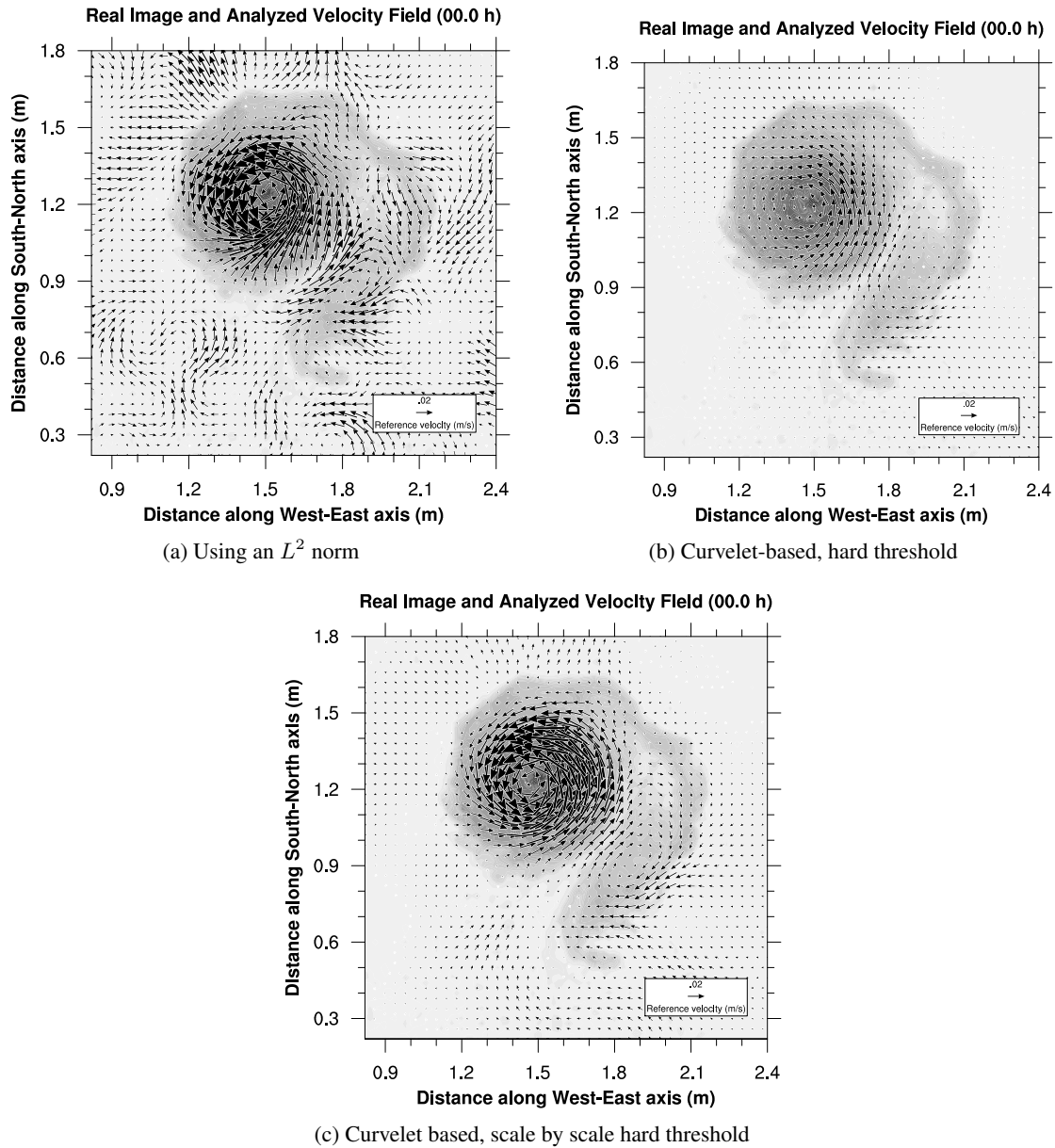


FIGURE 2.15: Initial velocity field reconstructed by direct assimilation of a real image sequence using different observation operators and norms.

Observation operators based on Lagrangian Coherent Structures

Using a Synthetic Image Sequence to define a Structure Observation may be difficult because numerical advection schemes are known to smooth discontinuities out. This drawback may lead to representativeness error: observation model may not be able to represent observed image patterns. To overcome this problem, [Titaud et al. \(2011\)](#) suggest to define observation operators based on the computation of Lagrangian coherent structures of the fluid.

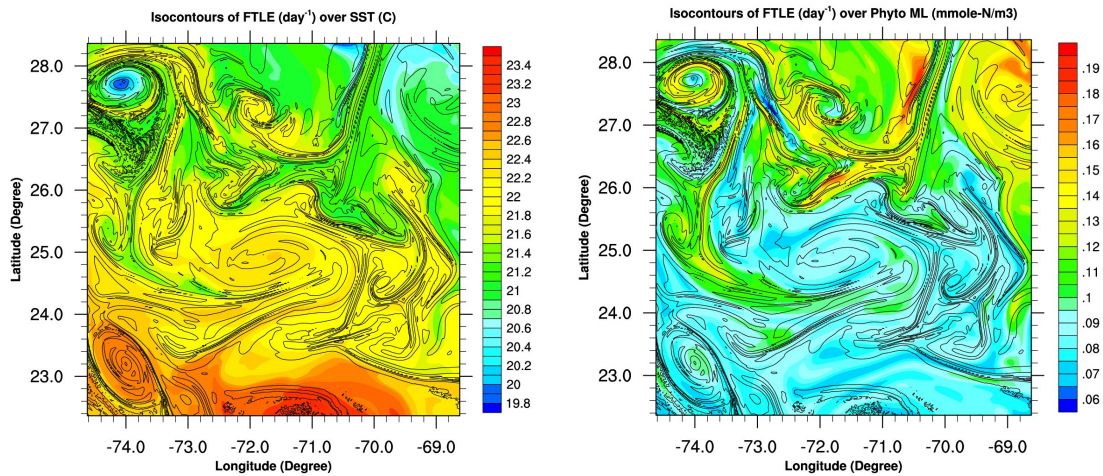
Lagrangian Coherent Structures (LCS) delimit regions of whirls, stretching, or contraction of tracer ([Ottino, 1989](#)). This concept is well defined for time-independent dynamic systems, where they simply correspond to stable and unstable manifolds of hyperbolic trajectories ([Wiggins, 1992](#)): contraction is observed along stable manifolds whereas unstable manifolds correspond to divergent directions along which the tracer is stretched. This concept was generalised for flows with general time dependence (particularly geophysical flows) ([Haller and Yuan, 2000](#); [Haller, 2011](#)) and they are usually identified in a practical manner as maximising ridges of so called backward Finite-Time Lyapunov Exponents (FTLE) field ([Haller, 2000, 2001, 2002](#); [Shadden et al., 2005](#); [Mathur et al., 2007](#)). Let $\mathbf{x}(t) = \mathbf{x}(t; \mathbf{x}_0, t_0)$ be the position of a Lagrangian particle at time t , started at \mathbf{x}_0 at $t = t_0$ and advected by the time-dependent fluid flow $\mathbf{u}(\mathbf{x}, t)$, $\mathbf{x} \in \Omega \subset \mathbb{R}^2$, $t \in [t_0, t_0 + T]$. The *forward* FTLE $\sigma_{t_0}^{t_0+T}(\mathbf{x}_0)$ at is defined as

$$\sigma_{t_0}^{t_0+T}(\mathbf{x}_0) = \frac{1}{|T|} \ln \sqrt{\lambda_{\max}(\Delta)}, \quad (2.14)$$

where λ_{\max} is the maximum eigenvalue of the Cauchy-Green strain tensor $\Delta = \left[\nabla \phi_{t_0}^{t_0+T}(\mathbf{x}_0) \right]^* \left[\nabla \phi_{t_0}^{t_0+T}(\mathbf{x}_0) \right]$, where $\phi_{t_0}^{t_0+T}$ is the flow map of the lagrangian system. FTLE corresponds to the growth factor of the norm of a perturbation started around \mathbf{x}_0 and transported by the flow after the finite advection time T . The associated eigenvector $\varphi_{t_0}^{t_0+T}(\mathbf{x}_0)$ is referred to as the *forward* Finite-Time Lyapunov Vector (FTLV): it corresponds to the direction of maximal stretching of the tracer. FTLE represents the rate of separation of initially neighbouring particles over the finite-time window $[t_0, t_0 + T]$. *Backward* FTLE-Vs are similarly defined, with the time direction being inverted in the computation of the Lagrangian trajectory. For more details on the practical computation of FTLE-V see e.g. [Shadden et al. \(2005, 2009\)](#) and [Ott \(1993\)](#) for any types of flows and [d’Ovidio et al. \(2004\)](#) for oceanic flows. FTLE (FTLV) is a scalar (vector) that is computed at a given location \mathbf{x}_0 . Seeding a domain with particules initially located on a grid leads to the computation of a discretised scalar (FTLE) and vector (FTLV) fields. Note that the FTLE-V naming may be misleading: they are more like practical ways of computing singular vectors of the transport equation rather than actual Lyapunov vectors (see chapter 3)

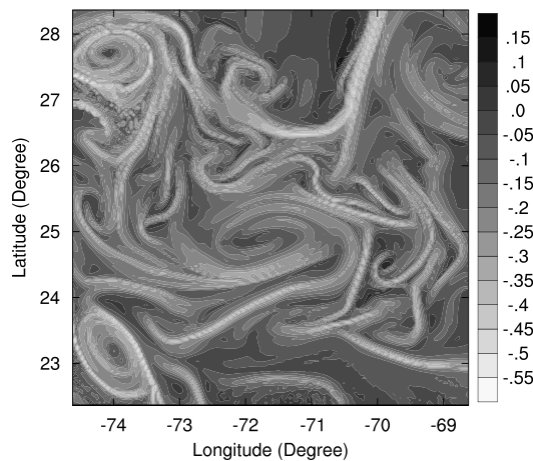
Backward FTLE fields show contours that correspond reasonably well to the main structures such as filaments, fronts and spirals that appear in geophysical and bio-geochemical tracer fields ([Beron-Vera and Olascoaga, 2009](#); [Shadden et al., 2009](#); [Olascoaga et al., 2008, 2006](#)). Figure 2.16 (top panels) shows the contours of the FTLE field (2.14) over Sea Surface Temperature (SST) (left panel) and Mixed Layer Phytoplankton (MLP) (right panel). The FTLE field is shown in the bottom panel.

FTLE are computed from a sequence of *mesoscale* ($1/4^\circ$) time-dependent velocity fields filtered from a high-resolution ($1/54^\circ$) idealised simulation of the North Atlantic Ocean (carried out by [Lévy et al. \(2009\)](#)). High resolution tracer fields SST and MLP come from the same simulation. Contours of FTLE matches quite well the main patterns in the corresponding SST and MLP fields of this simulation. [Lapeyre \(2002\)](#) shows that for a freely decaying 2D turbulence flow, the orientation of the gradient of a *passive* tracer converges to that of backward FTLVs. Such alignment properties have also been observed for realistic oceanic flows and tracers ([d’Ovidio et al., 2009b](#)). Figure 2.17 (bottom panel) shows the orientations of the backward



(a) Contours of FTLE over Sea Surface Temperature

(b) Contours of FTLE Phytoplankton



(c) Finite-Time Lyapunov Exponents field

FIGURE 2.16: Contours of Finite-Time Lyapunov Exponents over Sea Surface Temperature and Mixed Layer phytoplankton. From a $1/54^\circ$ North Atlantic Ocean simulation.

FTLV that correspond to the aforementioned FTLE contour field. Tracer gradients (left panel: SST; right panel: MLP) also show similar orientations. These two properties (pattern matching between tracer and FTLE scalar fields and tracer gradient orientation alignment with FTLV orientation) should be exploited in a Direct Image Assimilation framework. Furthermore, [Beron-Vera et al. \(2010\)](#); [Beron-Vera \(2010\)](#) showed, using real data, that these properties remain valid with a mesoscale advection, i.e. when the resolution of the velocity field — from which FTLE-V are computed — is much lower than the resolution of the observed tracer field. This behaviour was also mentioned using another Lagrangian tool (FSLE) by [Lehahn et al. \(2007\)](#); [d’Ovidio et al. \(2009b\)](#). Our synthetic data exhibit the same behaviour: FTLE field of Figure 2.16 is computed from a mesoscale $1/4^\circ$ velocity field on the same high-resolution ($1/54^\circ$) grid as that used to compute the SST and MLP fields. This feature is crucial from both a practical and physical point of view: first, velocity fields obtained from ocean global circulation models do not often provide more than mesoscale information, whereas tracer images contain submesoscale information; second, FTLE-V may be

used to quantify and characterise the link between scales.

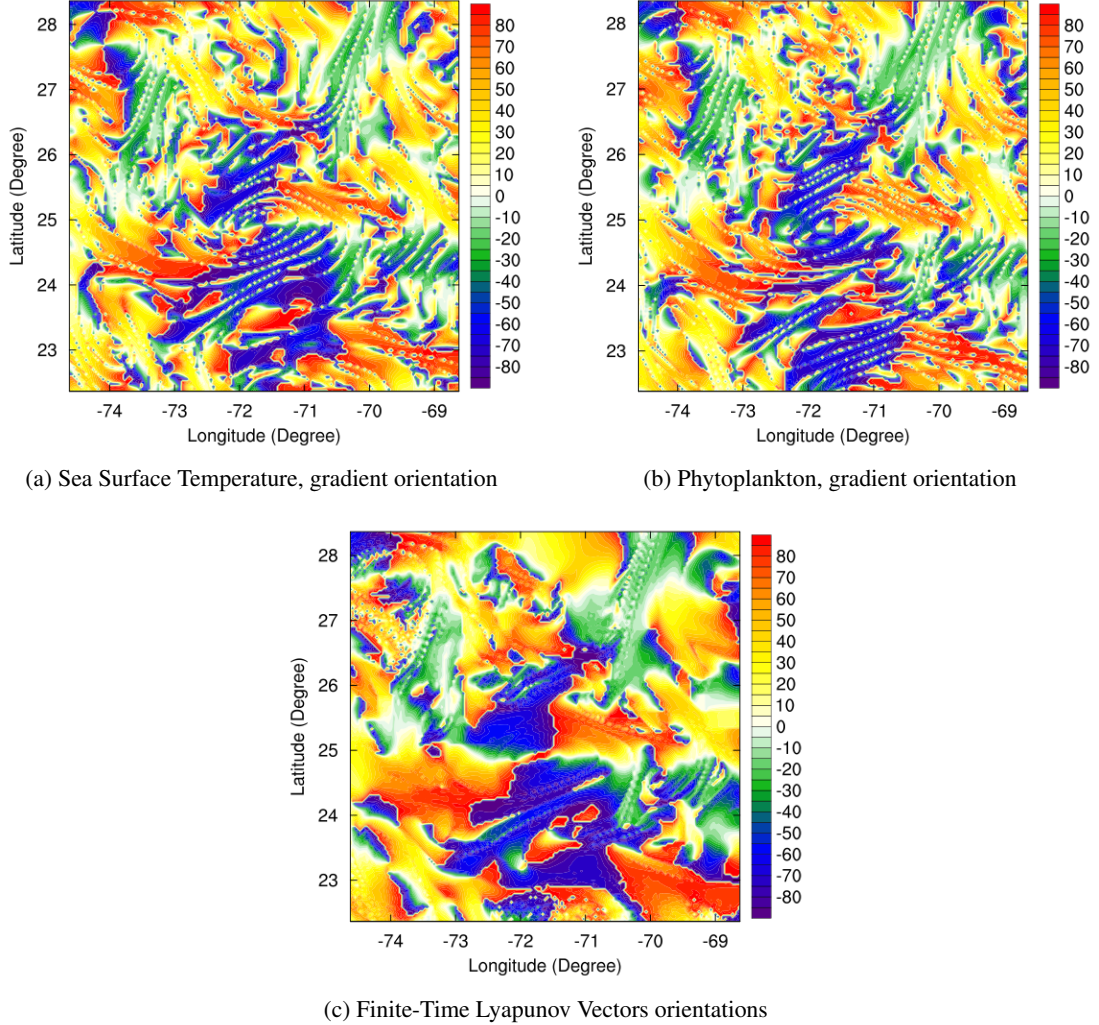


FIGURE 2.17: Orientations of Sea Surface Temperature and Mixed Layer phytoplankton gradient. Orientation of Finite-Time Lyapunov Vectors. From a $1/54^\circ$ North Atlantic Ocean simulation.

[Titaud et al. \(2011\)](#) suggest constructing observation operators using information given by the computation of Lagrangian Coherent Structures. The main idea is to exploit the two aforementioned properties of the FTLE-V to construct a triplet of direct assimilation of images:

- FTLE: the observation space \mathcal{I} is defined as a space of contours. Backward FTLE is viewed as an observation operator acting on the flow field \mathbf{u} (subset of control variable) onto the image space \mathcal{I} :

$$\begin{aligned} \mathcal{H}[\mathbf{u}] &: \Omega \rightarrow \mathbb{R} \\ \mathbf{x} &\mapsto \sigma_{t_0}^{t_0+T}(\mathbf{x}). \end{aligned} \quad (2.15)$$

- FTLV: the observation space is the space of function on \mathbb{R}^2 with values in the Euclidean unit sphere

equipped with an angular distance. The observation operator is simply the FTLV field:

$$\begin{aligned} \mathcal{H}[\mathbf{u}] &: \Omega \rightarrow \mathbb{R} \\ \mathbf{x} &\mapsto \varphi_{t_0}^{t_0+T}(\mathbf{x}). \end{aligned} \quad (2.16)$$

For the moment, few studies exist about the use of this kind of operator in direct image assimilation. In [Titau et al. \(2011\)](#) a sensibility study proves the potential of this tool by showing that computation of FTLE-V is sensitive (in the direct image assimilation framework) to small perturbations of the velocity field. [Figure 2.18](#) shows the variation of a sensitivity function (cost function) associated with the FTLE-V based observation operators with respect to the amplitude of nine random perturbations applied to a reference velocity field \mathbf{u} . Variations clearly have a minimum at the reference state and exhibit a convex shape that suggests a good situation for the inversion.

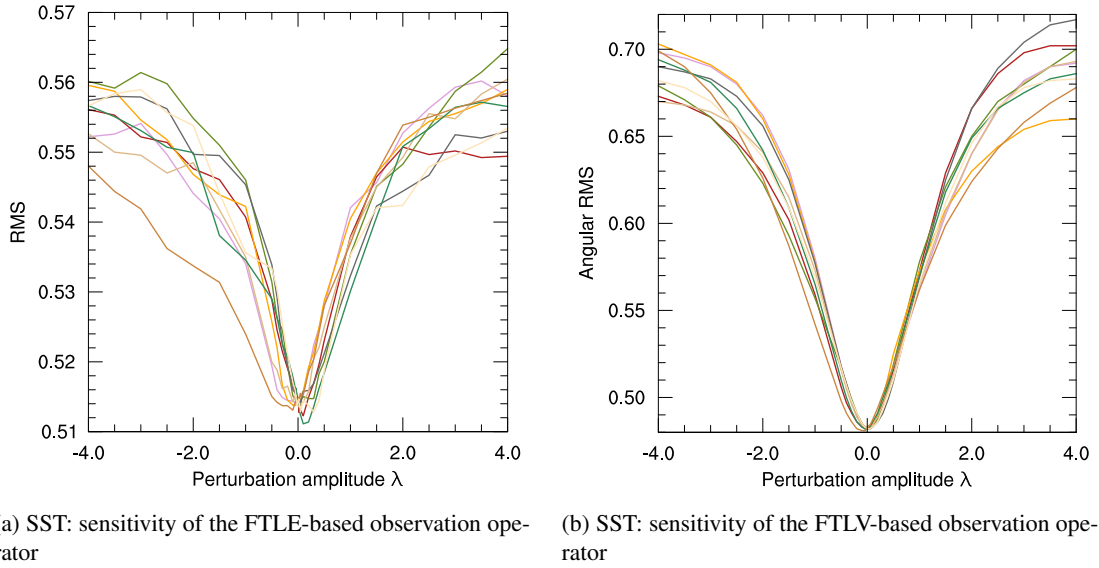


FIGURE 2.18: Variation of a sensitivity function (cost function) associated with the FTLE-V based observation operators with respect to the amplitude of nine random perturbations applied to a reference velocity field \mathbf{u} .

Even if not used in this description we also mention the Finite-Size Lyapunov Exponents (FSLE) ([Aurell et al., 1997](#); [Artale et al., 1997](#)) which is another Lagrangian tool that is commonly used in oceanographic contexts for studying mixing processes ([d'Ovidio et al., 2004, 2009a,b](#); [Lehahn et al., 2007](#), and references therein). In [Gaultier et al. \(2012\)](#) such FSLE field was used to reconstruct a ocean surface velocity field.

2.3 Conclusions et perspectives de ce chapitre

Comme on l'a vu au cours de ce chapitre, les problématiques soulevées par les différents types d'observation peuvent être très variées et l'assimilation de certaines d'entre elles peut se révéler très délicate, même pour des observations directes. On n'a cité ici que deux exemples, mais d'autres types d'observation présents ou futurs posent d'intéressants problèmes lorsqu'il s'agit de les assimiler. On peut ajouter à la variétés d'observations issues des satellites dont on a évoqué une partie dans ce chapitre, par exemple la

tomographie (Liu, 1995) qui fournit une observation intégrée en espace, ou l'assimilation d'estimations de transport à travers une section océanique (Stepanov et al., 2012), ou encore celle de positions lagrangiennes des flotteurs (Nodet, 2006, 2009; Vernieres et al., 2011) qui demandent d'ajouter un modèle de transport dans l'opérateur d'observation et/ou d'augmenter le vecteur de contrôle. Ce dernier type d'observation a d'ailleurs été inclus dans les système NEMOVAR dans le cadre de VODA (Chauvin et al., 2011). Cette liste n'est bien entendu pas exhaustive et nombre d'autres observations plus ou moins exotiques sont disponibles et ne demandent qu'à être assimilées.

Un aspect important, souvent ignoré par les assimilateur, et que l'on a évoqué à plusieurs reprise dans les deux premiers chapitre est la corrélation des erreurs d'observation entre elles. Si la présence de celles ci est évidente pour des observations de type satellite ou encore plus image, elle est également vraisemblable sur les données *in situ*. Cela peut-être du fait des instruments eu même, comme pour les profileurs XBT dont les incertitudes sur la vitesse de descente affectent l'ensemble des données de façon similaire. Une autre source est l'erreur de représentativité (le modèle ne peut représenter les phénomènes observés), qui potentiellement affecte tout type de donnée et peut représenter la majeure partie de l'erreur. Pour gérer cette dernière on augmente généralement la variance des erreurs d'observation de façon artificielle et dans le meilleur des cas en évaluant cette augmentation empiriquement. Richman et al. (2005) a proposé une approche plus systématique visant à projeter l'écart aux observation dans un sous espace dans lequel il n'y a plus d'erreur de représentativité. Ces résultats sont préliminaires, mais la piste semble être intéressante.

Outre ces aspects de représentation des erreurs d'observation, il me semble important de continuer à explorer la voie de l'assimilation d'image. Ici les pistes sont nombreuses et il est difficile de savoir laquelle ou lesquelles seront les plus prometteuses. Un gros travail sur la définition de l'espace image et de la distance associé est en cours dans notre équipe et ce dans plusieurs directions distinctes. Sans entrer dans les détails, on peut citer l'exploration d'autres décomposition multi-échelles comme les ondelettes, par exemple, l'utilisation de curvelettes 3D (2D+temps) afin d'obtenir une meilleurs cohérence temporelle Ma et al. (2008); Toye Mahamadou Kele (2009), des distances basées sur l'orientation des gradients de l'image (Papadakis and Chabot, en préparation) ou basées sur du transport optimal... beaucoup de travail en perspective.

Chapitre 3

DES MODÈLES INCOMPRIS

Si dans un système de prévision il sera toujours nécessaire de procéder à de l'assimilation de données pour prévenir la dérive des modèles (processus ignorés, conditions aux limites mal connus, ...), il ne doit pas être considéré que cela est suffisant pour obtenir une bonne prévision et il est important de se consacrer à l'amélioration des modèles (et pas seulement en augmentant la résolution). Pour cela l'assimilation de données ou les outils issus de celle-ci peuvent être d'une grande aide. Ces outils, même les plus simples sont assez peu utilisés pour l'océanographie de façon systématique (c'est moins vrai pour l'atmosphère). De mon point de vue, cette partie est plutôt prospective avec toutefois quelques résultats préliminaires sur les aspects les plus simples. Cette prospective est amplement motivée par les efforts que j'ai déployé ces dernières années, notamment durant le projet VODA, afin de bâtir les modèles tangent et adjoint (TAM) de NEMO tout en proposant une stratégie pour pérenniser ces outils (Vidard et al., 2009b, 2012). L'objectif initial de ces travaux était de participer au développement de NEMOVAR une plateforme d'assimilation variationnelle pour NEMO. Cet objectif a été atteint, NEMOVAR étant devenu le système opérationnel des analyse océaniques à l'ECMWF et au Met-Office du Royaume-Uni (dans sa version 3D-Fgat) et que NEMOTAM est désormais mis à disposition de la communauté scientifique. Cependant, un argument pour soutenir ce pour développement était de mettre en évidence les multiples autres applications possibles des TAM.

Dans ce chapitre on se propose donc de présenter un panorama de ces méthodes qui, à l'interface des méthodes d'assimilation de données permettent une meilleur compréhension du fonctionnement des modèles et devraient permettre leur amélioration et/ou leur validation. Certaines sont déjà connue, bien que peu utilisées pour les problèmes d'océan ou d'atmosphère, d'autres proviennent de domaines applicatifs bien différents et doivent donc être adaptés aux spécificités des problèmes qui nous intéressent.

3.1 Étude de l'incrément d'analyse

Un des premiers enseignements sur le comportement d'un modèle que l'on peut obtenir directement grâce aux outils d'assimilation peut provenir de l'étude des séries temporelles des incréments d'analyse eux même. En effet l'existence de structures récurrentes dans ces séries (en gros on corrige toujours la même chose au même endroit, dans la même direction) peut mettre en évidence des dysfonctionnements du modèle. Il faut cependant prendre garde à ne pas tirer de conclusions trop hâtives, ces artefacts pouvant être générés par l'assimilation elle même. Une belle illustration simultanée de ces deux aspects peut être trouvé dans nos travaux sur l'élaboration du nouveau système d'analyse océanique de l'ECMWF (Balmaseda et al., 2008) dont est tiré la figure 3.1. Celle ci présente le même diagnostic sur l'ancien (ORA-S2) et le

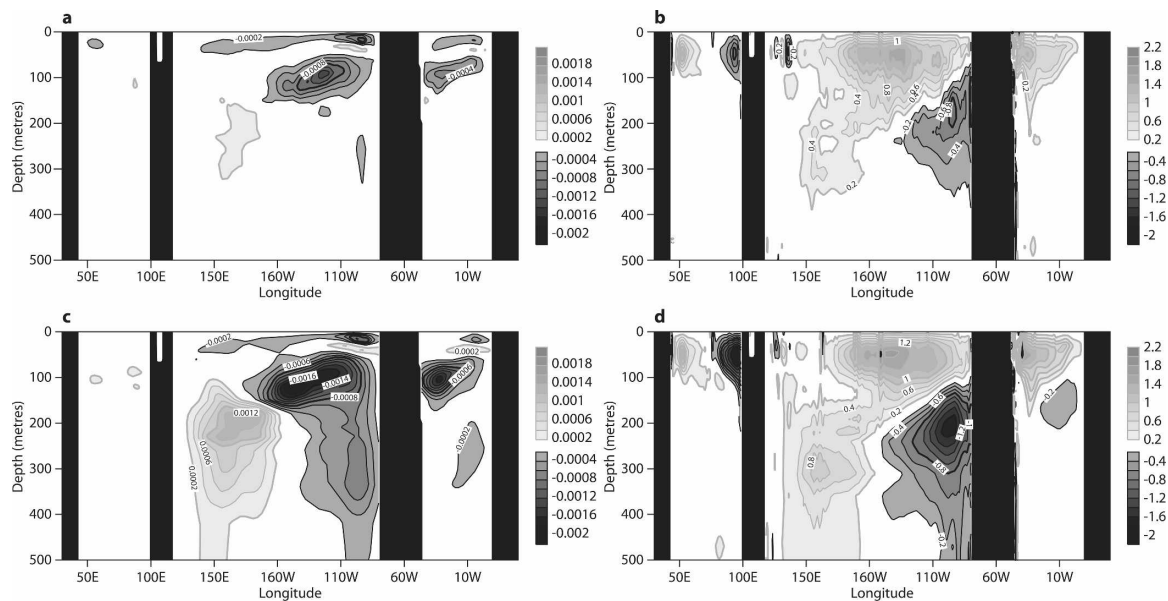


FIGURE 3.1: Section équatoriale longitude-profondeur de (a),(c) l'incrément moyen de température et (b),(d) la vitesse verticale pour les systèmes ORA-S2 et ORA-S3 respectivement sur la période 1987-2001.

nouveau (ORA-S3) système d'assimilation de données océanique, le modèle étant le même dans les deux cas (HOPE). Le nouveau système produit (heureusement) de meilleurs résultats notamment en terme de proximité avec les données alors que la comparaison des panneaux (a)-(b) avec (c)-(d) montre que l'incrément moyen est plus petit. Le modèle étant le même ceci est plutôt bon signe et symptomatique d'un problème dans le précédent schéma d'assimilation. Cette déficience est mise en évidence par l'incrément positif en température dans le Pacifique Ouest (en gris clair sur le panneau (c)). Celui-ci est associé à une circulation verticale indésirable qui était provoqué par un déséquilibre dans l'incrément calculé par ORA-S2. La circulation induite par ce déséquilibre amenait des eaux des niveaux inférieurs plus froide en surface, le système d'assimilation réagissait alors en réchauffant cette zone pour mieux correspondre aux données. L'augmentation de l'incrément entraînait une circulation verticale non désirée encore plus grande et ainsi de suite. En revanche l'incrément négatif résiduel dans le Pacifique Est (panneau (a)) est lié à une déficience connue du modèle dans le resserrement de la thermocline. Ce genre de diagnostic basique, produit direct du système d'assimilation, ne permet pas de différencier les défaillances du modèle ou du schéma d'assimilation, mais permet au moins de se poser de bonnes questions et donne des pistes d'investigation pour améliorer l'ensemble du système de prévision. Ces pistes pourront ensuite être explorées par des approches plus évoluées comme celles présentées dans les sections suivantes.

3.2 Analyse de sensibilité

Une des approches à fort potentiel pour l'aide aux modélisateurs est ce que l'on regroupe sous le terme d'Analyse de Sensibilité. Il s'agit ici de quantifier comment une ou des sorties données d'un système donné est sensible à des variations sur une ou des entrées données. Extraire ce genre d'information peut avoir plusieurs intérêts : (i) comprendre le système en mettant en évidence les paramètres importants; (ii) déterminer les effets sur le comportement du système des variations de ces paramètres; (iii) construire et optimiser

ce système; (iv) focaliser les améliorations du système sur les parties les plus sensibles, Cela permet également, et c'est moins mis en évidence dans la littérature, de valider un modèle. En effet une sensibilité inattendue devra permettre de remettre en cause soit le modèle lui-même, soit les connaissances a priori qu'on a sur le fonctionnement réel du système.

Il existe de nombreuses façons de faire de l'analyse de sensibilité, la plus basique étant bien entendu de choisir un paramètre et de faire plusieurs simulations identiques en ne faisant varier que ce paramètre. Si cette approche est peu coûteuse et très facile à mettre en place, on comprendra aisément qu'elle ne permet pas de conclure en raison de la faible exploration de l'espace du paramètre. Cette pratique a été améliorée pour créer ce qu'on appelle les méthodes OAT (One At a Time) qui conservent les estimations de chaque paramètre individuellement mais proposent une exploration plus systématique de l'espace des paramètres, comme c'est le cas de la méthode de [Morris \(1991\)](#).

On peut classer les méthodes d'analyse de sensibilité en deux grandes classes : les méthodes globales et les méthodes locales. Les premières étant en générale stochastiques et les secondes déterministes. On se retrouve donc, comme c'était le cas pour l'assimilation de données, à l'interface de ces deux aspects. Cependant, et contrairement au cas de l'assimilation de données, la nature de l'information obtenue est différente dans les deux cas.

Dans la suite de cette section, on suppose qu'on dispose d'un modèle d'évolution d'un système

$$\begin{cases} \frac{d\mathbf{x}}{dt} &= \mathcal{M}(\mathbf{x}, \boldsymbol{\alpha}), \quad t \in [0, T], \\ \mathbf{x}(0) &= \mathbf{x}_0, \end{cases}$$

avec un état $\mathbf{x} = (x_1, \dots, x_n)^T \in \mathcal{X} \subset \mathbb{R}^n$ et jeu de paramètres $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_p)^T \in \mathcal{P} \subset \mathbb{R}^p$ et on définit une fonction réponse (ou coût, ou objective, ou critère) qui représente la quantité sensible

$$\begin{aligned} J : \quad \mathcal{P} &\rightarrow \mathbb{R} \\ \boldsymbol{\alpha} &\mapsto J(\boldsymbol{\alpha}) = \int_0^T \phi(\mathbf{x}, \boldsymbol{\alpha}, t) dt \end{aligned}$$

où ϕ fonctionnelle non linéaire à valeur dans \mathbb{R} . On a considéré ici le cas mono-critères (J à valeurs dans \mathbb{R}) mais ce qui suit est facilement généralisable au cas multi-critère.

3.2.1 Analyse globale

L'analyse de sensibilité globale étudie comment la variabilité des entrées se répercute sur celle de la sortie, généralement en déterminant quelle part de variance de la sortie est due à telle entrée ou tel ensemble d'entrées. Si l'analyse de sensibilité locale s'intéresse plus à la valeur de la réponse, l'analyse de sensibilité globale s'intéresse quant à elle à sa variabilité.

L'Analyse de sensibilité globale est généralement effectuée par approche stochastique dont l'objectif est des déterminer un critère objectif basée sur les densités de probabilités des entrées et sorties d'un modèle déterministe. Pour les méthodes basées sur la variance, on définit l'indice de sensibilité au premier ordre par :

$$S_i = \frac{\text{Var}_{\alpha_i} [\mathbb{E}(J|\alpha_i)]}{\text{Var} [\mathbb{E}(J)]} \quad (3.1)$$

où $\mathbb{E}(\cdot)$ représente l'espérance mathématique et $\text{Var}(\cdot)$ la variance.

Cet indice mesure l'effet principal de α_i sur J (*i.e.* la fraction de la variance de $J(\alpha)$ due à la contribution de α_i). On peut généraliser ce concept en définissant les indices de sensibilité d'ordre k si on pose $\mathcal{I} = \{i_1, \dots, i_k\} \subset \{1, \dots, p\}$ et $\alpha_{\mathcal{I}} = (\alpha_{i_1}, \dots, \alpha_{i_k})$ et

$$S_{\mathcal{I}} = \frac{\text{Var}_{\alpha_{\mathcal{I}}} [\mathbb{E}(J|\alpha_{\mathcal{I}})]}{\text{Var} [\mathbb{E}(J)]} - \sum_{I \subset \mathcal{I}, I \neq \mathcal{I}} S_I \quad (3.2)$$

Ces derniers peuvent être considéré comme des mesures des effets d'interaction, c'est à dire la part de la variation de J due aux éléments de $\alpha_{\mathcal{I}}$ qui ne peut pas être expliquée par la somme des effets individuels de ses éléments (ou sous ensembles d'élément). On voit bien que le nombre d'indice à calculer devient rapidement très importants à mesure que le nombre p de paramètres augmente (on calcule $2^p - 1$ indices). [Homma and Saltelli \(1996\)](#) ont alors introduit la sensibilité totale du système au $i^{\text{ème}}$ facteur :

$$S_i^T = \sum_{\mathcal{I} \subset \{1, \dots, p\}, i \in \mathcal{I}} S_{\mathcal{I}} \quad (3.3)$$

On peut montrer facilement que ces indices peuvent être estimés directement (sans passer par l'estimation des composantes de la somme) par :

$$S_i^T = 1 - \frac{\text{Var}_{\alpha_{\alpha-i}} [\mathbb{E}(J|\alpha_{\alpha-i})]}{\text{Var} [\mathbb{E}(J)]} \quad (3.4)$$

où $\alpha_{\alpha-i}$ représente toutes les composantes de α sauf α_i . Le calcul de ces indices se fait le plus souvent par une décomposition de la surface de réponse (décomposition de Fourier pour la méthode FAST ([Saltelli et al., 2000](#)), décomposition fonctionnelle de type ANOVA (ANalysis Of VAriance) pour la méthode de Sobol' ([Sobol, 1990](#))).

Le calcul des indices de Sobol' requiert l'évaluation de nombre d'intégrales de $J(\alpha)$; ce nombre augmente avec la taille du problème. On utilise classiquement des méthodes de type Monte-Carlo, ce qui requiert généralement un nombre N d'estimation de J de l'ordre de quelques centaines. Suivant le coût d'un calcul de J , le coût de l'estimation des indices de Sobol', même pour les seuls S_i et S_i^T peut devenir rapidement prohibitif (cela requiert $N(2p + 2)$ évaluation de J). De plus, et c'est sans doute plus problématique, les méthodes de Sobol' ou FAST reposent sur l'hypothèse que les paramètres sont décorrélés ce qui limite l'applicabilité de telle méthodes à des problèmes bien particuliers. Appliquer ce type de méthodes à des problèmes du type de ceux qui nous intéresse ici est l'objet de nombreuses recherches actuellement (voir par exemple ([Chastaing et al., 2012](#)) et les références incluses pour un récent tour d'horizon).

Pour ces raisons l'analyse de sensibilité globale a encore été très peu appliquée à des problèmes en océanographie, les quelques applications que j'ai pu trouver le sont à la biologie où typiquement une dizaine de paramètre indépendants et mal connus peuvent être isolés (confère par exemple [Chu et al. \(2007\)](#))

3.2.2 Analyse locale

L'analyse de sensibilité locale consiste en calculer de façon exacte et efficace les sensibilités de la réponse du système à des variations de ses paramètres autour leur valeur initiale (d'où la localité).

En d'autres termes, on veut savoir quelle est la sensibilité de J aux variations des éléments de α , ce qui se formalise par trouver

$$S_{\alpha} = \frac{dJ}{d\alpha}(\alpha) = \nabla_{\alpha} J(\alpha) = \left(\frac{\partial J}{\partial \alpha_1}(\alpha), \dots, \frac{\partial J}{\partial \alpha_p}(\alpha) \right)^T \quad (3.5)$$

S_α est le vecteur des sensibilité locale de J aux variations de α , c'est une sensibilité locale car dépendante de l'état actuel de α (et de \mathbf{x}_0).

Le plus simple est évidemment d'approcher S_α en calculant les taux d'accroissement

$$S_\alpha^i \approx \frac{J(\alpha + (0, \dots, \delta\alpha_i, \dots, 0)^T) - J(\alpha)}{\delta\alpha_i} \quad (3.6)$$

pour tout i . Par contre le coût de calcul augmente rapidement avec la dimension de α et le choix de $\delta\alpha_i$ est critique (trop grand l'erreur de troncation devient importante, trop petit ce sont les erreurs d'arrondis)

Pour pallier ce dernier point, on peut obtenir un calcul exact par les dérivées de Gâteaux (FSAP, *Forward Sensitivity Analysis Procedure* dans la terminologie de [Cacuci \(2003\)](#)):

$$\hat{J}(\bar{\alpha})[\hat{\alpha}] = \int_0^T \left[\frac{\partial\phi}{\partial\mathbf{x}} \right]_{\bar{\alpha}} \hat{\mathbf{x}} + \left[\frac{\partial\phi}{\partial\alpha} \right]_{\bar{\alpha}} \hat{\alpha} dt \quad (3.7)$$

avec $\hat{\mathbf{x}}$ la dérivée de Gâteaux de \mathbf{x} dont on connaît l'évolution en appliquant la définition sur le modèle

$$\begin{cases} \frac{d\hat{\mathbf{x}}}{dt} &= \left[\frac{\partial\mathcal{M}}{\partial\mathbf{x}} \right]_{\bar{\alpha}} \hat{\mathbf{x}} + \left[\frac{\partial\mathcal{M}}{\partial\alpha} \right]_{\bar{\alpha}} \hat{\alpha}, \quad t \in [0, T], \\ \hat{\mathbf{x}}(0) &= \hat{\mathbf{x}}_0, \end{cases} \quad (3.8)$$

où $\left[\frac{\partial\mathcal{M}}{\partial\mathbf{x}} \right]_{\bar{\alpha}}$ est le Jacobien du modèle par rapport à la variable d'état. Ce système est appelé Modèle Linéaire Tangent. Il permet de calculer de façon exacte la dérivée directionnelle de la fonction coût. Cependant, la résolution du système linéaire tangent est effectuée pour un $\hat{\mathbf{x}}$ donné, déterminé par une direction $\hat{\alpha}$ donnée. Pour obtenir toutes les composantes de S_α il faut calculer ces dérivées directionnelles dans la direction de chaque éléments de la base de \mathcal{P} . On a donc résolu le problème de précision, mais pour un nombre de paramètre important le coût de calcul reste prohibitif.

On peut cependant utiliser, comme en assimilation variationnelle, une méthode adjointe pour calculer ce gradient (ASAP, *Adjoint Sensitivity Analysis Procedure* dans la terminologie de [Cacuci \(2003\)](#)), en posant le Modèle Adjoint (qui s'intègre de façon rétrograde) :

$$\begin{cases} \frac{d\mathbf{p}}{dt} + \left[\frac{\partial\mathcal{M}}{\partial\mathbf{x}} \right]_{\bar{\alpha}}^T \mathbf{p} &= \left[\frac{\partial\phi}{\partial\mathbf{x}} \right]_{\bar{\alpha}}^T, \quad t \in [0, T], \\ \mathbf{p}(T) &= 0 \end{cases} \quad (3.9)$$

on obtient de façon classique:

$$\nabla_\alpha J = \int_0^T \left[\frac{\partial\phi}{\partial\alpha} \right]_{\bar{\alpha}}^T - \left[\frac{\partial\mathcal{M}}{\partial\alpha} \right]_{\bar{\alpha}}^T \mathbf{p} dt \quad (3.10)$$

où $\nabla_\alpha J$ ne dépend plus des dérivées de Gâteaux $\hat{\mathbf{x}}$ et $\hat{\alpha}$. Ainsi le coût de calcul ne dépend plus du nombre de paramètre, et S_α peut être calculé de façon exacte au moyen d'une intégration du modèle direct et une intégration rétrograde du modèle adjoint (si le modèle direct est linéaire, seule l'intégration de l'adjoint sera nécessaire).

Cependant l'information obtenue par cette approche est nettement moins riche que celle fournie par analyse globale. On peut toutefois noter que dans le cas d'un modèle linéaire les S_i sont des indices globaux car il n'y a pas de termes d'ordre supérieurs dans le développement de Taylor et ils sont proportionnels à ceux de l'équation 3.1.

On peut trouver un exemple d'application de ce type de méthode dans l'étude que nous avons menée avec É. Rémy de Mercator-Océan sur les réanalyses GLORYS de l'océan global au $1/4^\circ$. L'objectif initial est de tenter d'évaluer les zones d'influences géographiques permettant de réduire l'erreur de prévision en utilisant une méthode adjointe pour calculer cette sensibilité (alors que le système d'assimilation de GLORYS est basé sur une méthode d'interpolation optimale). Nous avons effectué une étude préliminaire en considérant l'écart aux observations comme étant un proxy de cette erreur de prévision et en voulant déterminer la sensibilité de cet écart à des modifications de la condition initiale et/ou des forçages. Cela devrait nous donner une indication sur les phénomènes importants à considérer pour améliorer ce système.

L'exemple le plus facile à interpréter de cette étude est de considérer l'écart aux cartes de SST à l'instant final d'un cycle d'assimilation comme critère de sensibilité, en raison de sa couverture dense en espace. C'est à dire qu'on s'intéresse au calcul du gradient de :

$$J(\mathbf{x}_0, \mathbf{q}) = \frac{1}{2} \sum_{n=1}^{N_{SST}} \| H_{SST}(\mathbf{x}_n) - SST^{obs} \|_{\mathbf{R}^{-1}}^2 \quad (3.11)$$

avec un vecteur de contrôle constitué de $\mathbf{x}_0 = (u_0, v_0, T_0, S_0, \eta_0)^T$ le vecteur d'état initial (composantes de la vitesse de courant, température, salinité et surface libre) et de $\mathbf{q} = (q_{sr}, q_{ns}, emp)^T$ (flux radiatif, flux de chaleur total, flux d'eau douce).

Sans rentrer dans les détails de cette étude qui n'a pas encore été publiée, on peut regarder l'exemple des sensibilité à la température initiale (en surface et à 100m) comme montré sur les deux panneaux du bas de la figure 3.2. Une forte sensibilité donnera un signal ressemblant à l'écart aux observations (en haut à gauche), tandis qu'une faible sensibilité donnera une carte blanche. Dans cet exemple on voit bien que l'écart à la SST est fortement sensible aux modifications de température de surface initiale où la profondeur de couche de mélange (en haut à droite) est faible et peu sensible ailleurs. Le constat inverse peut être fait pour la sensibilité à la température initiale à 100m. Ceci n'est évidemment pas une surprise, et correspond plus à l'objectif de vérification du modèle (et de l'approche d'A.S.) qu'à celui original d'amélioration du système d'assimilation. Cependant cela met en lumière l'importance d'avoir une bonne estimation du mélange vertical et fait écho aux problèmes de déstabilisation de ce mélange rencontrés au début de ce chapitre. D'autres composantes du gradient montrent le rôle important des forçages atmosphériques (là encore on s'en serait douté) et les pistes pour l'amélioration du système semblent aussi se diriger dans cette voie. Dans un objectif d'amélioration du système d'assimilation de données, cette approche n'est évidemment pas complètement satisfaisante ; en effet en toute rigueur il conviendrait d'inclure le système d'assimilation dans le système d'optimalité en en contrôlant son ébauche. En théorie, ce système d'assimilation étant linéaire et matriciel, en dériver son adjoint devrait être chose aisée, en pratique c'est une toute autre histoire, manipuler un système opérationnel n'est jamais chose aisée, mais le jeu en vaut sans doute la chandelle.

3.2.3 De l'analyse locale à l'analyse globale

Dans la section précédente on a rapidement présenté les méthodes classiques d'analyse de sensibilité par calcul (ou approximation) de gradient. Ces analyses sont locales au sens où elles dépendent de l'état autour duquel on a dérivé notre fonction réponse. Ainsi ces analyses peuvent donner des sensibilités significativement différentes suivant l'état de référence.

Il y a plusieurs approches possibles afin de "globaliser" l'analyse par calcul de gradient. [Cacuci \(2003\)](#) joue sur la définition d'analyse globale et propose comme nouvel objectif pour celle-ci de rechercher les points critiques de la fonction réponse dans l'espace des paramètres (bifurcations, points d'inflexion, extrema, ...) et ensuite d'effectuer une analyse locale de ces points critiques. Dans les cas qui nous intéressent

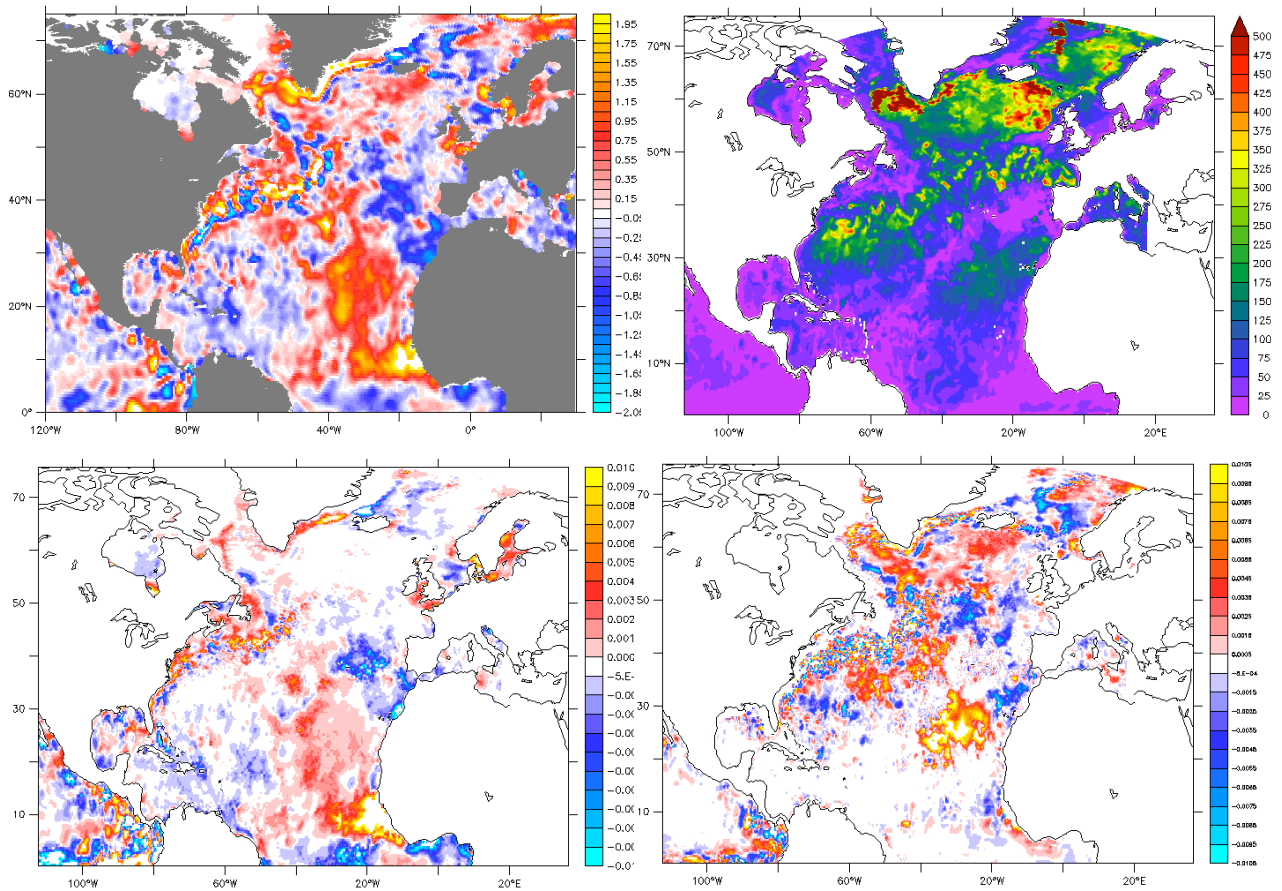


FIGURE 3.2: Haut : Ecart entre la SST prévue et observée (gauche) et profondeur de la couche de mélange (droite) Bas : sensibilité de l'erreur en SST après une semaine de prévision par rapport à des variations de température initiale en surface (gauche) et à 100m (droite)

leur calcul est loin d'être évident, les extremums peuvent être trouvés par optimisation, tandis que les points d'inflexions demanderont une analyse au second ordre.

Une autre approche consiste, à l'instar de ce qui se fait en contrôle robuste (cf plus loin) à globaliser l'indice de sensibilité en en prenant l'espérance mathématique dans l'espace des paramètres :

$$M_i = \mathbb{E}(J) \quad (3.12)$$

En utilisant cette approche, [Sobol and Kucherenko \(2009\)](#) a montré que $\tilde{M}_i = \mathbb{E}(J^2)$ était comparable aux S_i^T de l'analyse globale basées sur la variance. De plus le calcul des indices basés sur les dérivées par méthodes de type Monte-Carlo ou Quasi Monte-Carlo sont moins coûteux que ceux de Sobol' de quelques ordres de grandeur ([Kucherenko et al., 2009](#)).

3.2.4 Propagation d'incertitude

Les systèmes de modélisation se développant et prenant une place de plus en plus importante dans nos sociétés, on assiste à un besoin croissant de barre d'erreur sur les prévisions. Les sensibilités calculées dans les sections précédentes peuvent être d'une grande utilité à cette fin. Si on considère que α est un vecteur

aléatoire dont on connaît la moyenne et la variance, on peut estimer ces quantités en sortie du système par propagation des moments. Cela revient à écrire le développement de Taylor à l'ordre 1 de J autour de $\bar{\alpha} = \mathbb{E}(\alpha)$

$$J(\alpha) \simeq J(\bar{\alpha}) + \sum_{i=1}^p \frac{\partial J}{\partial \alpha_i} \Big|_{\bar{\alpha}} (\alpha_i - \bar{\alpha}_i) = J(\bar{\alpha}) + \sum_{i=1}^p S_{\bar{\alpha}}^i (\alpha_i - \bar{\alpha}_i) \quad (3.13)$$

On peut alors facilement calculer l'espérance $\mathbb{E}(J) \simeq J(\bar{\alpha})$ et la variance

$$\text{Var}(J) = \mathbb{E}((J - \mathbb{E}(J))^2) \simeq \mathbb{E}\left(\left(\sum_{i=1}^p S_{\bar{\alpha}}^i (\alpha_i - \bar{\alpha}_i)\right)^2\right) = \sum_{i=1}^p S_{\bar{\alpha}}^i{}^2 \text{Var}(\alpha_i) + 2 \sum_{i \neq j=1}^p S_{\bar{\alpha}}^i S_{\bar{\alpha}}^j \text{Cov}(\alpha_i, \alpha_j) \quad (3.14)$$

On peut réécrire cette équation sous forme matricielle pour obtenir la *sandwich rule* :

$$\text{Var}(J) \simeq \mathbf{S} \mathbf{C}_{\alpha} \mathbf{S}^T \quad (3.15)$$

où \mathbf{S} est le vecteurs des sensibilités d'ordre 1 et \mathbf{C}_{α} la matrice de covariance de α .

On peut bien évidemment garder des termes d'ordre supérieurs dans le développement de Taylor, mais dans ce cas il faudra connaître les dérivées de J et surtout les moments de α d'ordre supérieurs (en gardant les termes d'ordre 2 on aura besoin des moments d'ordre 4 de α pour le calcul de la variance de J). Pour cette raison on se cantonne généralement aux calculs présentés ci dessus, ce qui se ramène à ne considérer, une fois de plus, que la partie linéaire de la propagation d'incertitude. Il est aussi important de noter que là encore si les éléments de α sont décorrélés cela simplifie grandement le problème, en effet le deuxième terme de l'expression de $\text{Var}(J)$ disparaît et la conservation des termes d'ordre 2 devient faisable en pratique.

Un gros avantage d'utiliser la méthode de propagation des moments est qu'on peut déduire facilement l'impact d'une amélioration de l'incertitude des entrées du système sur les sorties de ce même système si on dispose des sensibilités locales du premier ordre. L'inconvénient est bien sur, qu'il faut disposer de ces sensibilités, mais on a montré dans les sections précédentes qu'on pouvait en disposer à moindre coût grâce aux méthodes adjointes.

3.3 Estimation de paramètres et Contrôle robuste

Comme beaucoup de système de modélisation, l'océan est caractérisé par un certain nombre de paramétrisations ajustables dont les réglages optimaux sont mal connus (processus sous mailles, paramétrisation du mélange vertical, interaction avec l'atmosphère, ...). On peut bien évidemment améliorer le réglage de ces paramètres par contrôle optimal ou par filtrage stochastique pour peu qu'on soit à même de définir des critères adéquats à optimiser.

Bien qu'ils utilisent généralement les même méthodes, j'aurai tendance à différencier l'assimilation de données et le contrôle de paramètre, en effet dans le premier cas on cherche des conditions initiales ou aux limites optimales à un instant donné, alors que dans le deuxième cas on aimerait estimer un jeu de paramètre valide quelque soit la situation océanique. De plus dans le cas du contrôle de paramètres, le critère à optimiser n'est pas forcément un écart aux observations.

Cependant la valeur de la fonctionnelle J peut aussi dépendre d'autres facteurs $\xi = (\xi_1, \dots, \xi_m)^T \in \mathcal{A} \subset \mathbb{R}^m$ sur lesquels pèsent des incertitudes (typiquement les forçages atmosphériques en océanographie). On peut vouloir prendre en compte ces incertitudes dans l'estimation optimale de nos paramètres. Pour ce faire on peut utiliser ce que certaines communautés appellent le contrôle robuste. Ce type de méthode est principalement utilisé dans les secteurs à risque (sûreté nucléaire) et dans la finance.

Pour garder un cadre similaire à ce qui précède, on note $J = J(\alpha, \xi)$ la fonction objective qui devient du même coup une variable aléatoire. La valeur de J n'est plus univoque et chercher à l'optimiser peut avoir plusieurs sens. Sous l'appellation contrôle robuste plusieurs branches se retrouvent (qui peuvent avoir des noms distincts suivant les domaines d'application). La plus classique, mais peut-être la moins adaptée au contrôle en océanographie, est ce qu'on pourrait appeler le scénario du pire ou il s'agit de trouver un contrôle qui permet d'optimiser le système dans les plus mauvaises conditions sur la variable incertaine, c'est à dire trouver le point-selle de J ou le minimum de

$$\tilde{J}(\alpha) = \max_{\xi \in \mathcal{D}_\xi} J(\alpha, \xi) \quad (3.16)$$

Sans doute plus adapté, et en lien avec ce qui a été décrit à la fin de la section précédente, afin de prendre en compte ces incertitudes de façon moyenne, on peut alors considérer la minimisation d'une nouvelle fonctionnelle

$$\bar{J}(\alpha) = \mathbb{E}_\xi(J(\alpha, \xi)) = \int_{\mathcal{D}_\xi} J(\alpha, \xi) p(\xi) d\xi \quad (3.17)$$

où \mathbb{E}_ξ est l'espérance mathématique par rapport au paramètre incertain ξ , et $p(\xi)$ est la densité de probabilité de ξ .

Une telle approche fournit évidemment a priori une estimation plus robuste vis à vis des incertitudes de modélisation que l'estimation fournie par l'approche d'inversion classique décrite précédemment (qui minimise en fait quelque chose comme $J(\alpha, \mathbb{E}(\xi))$).

La robustesse recherchée peut aussi être une stabilité par rapport aux variations de la fonction objective, dans ce cas on aimerait contrôler la variance :

$$\check{J}(\alpha) = \text{Var}_\xi(J(\alpha, \xi)) \quad (3.18)$$

et ainsi de suite. La littérature sur le contrôle robuste est assez fournie, même si les domaines d'application sont assez lointains de celui qui nous intéresse ici, et nécessite donc une adaptation. Le choix de la robustesse recherchée dépend bien entendu du problème étudié, cependant pour nombre d'applications en océanographie il me semble qu'une combinaison des fonctionnelles 3.17 et 3.18 pourrait être un bon candidat, c'est à dire qu'on chercherait un contrôle optimal en valeur et en variabilité de la réponse. On chercherait alors à minimiser une fonctionnelle ressemblant à

$$\check{J} = \bar{J} + \gamma \check{J}. \quad (3.19)$$

Bien évidemment, comme c'était le cas dans l'analyse de sensibilité, la question du calcul de ces contrôles robustes se pose. Pour le calcul de \check{J} on peut, à l'instar de ce qui a été proposé pour la propagation d'incertitude, de se baser sur une approximation de Taylor pour calculer l'espérance et la variance (équation 3.14). Il est important de noter que dans ce cas il sera nécessaire d'avoir les dérivées d'ordre 2 de J (ou d'ordre 3 si on conserve l'ordre 2 dans le développement de Taylor avec des ξ_i décorrélés). La plupart des non-linéarités incluses dans les modèles d'océan étant de type quadratique, ce dernier point n'est pas si difficile à résoudre et cela limite les développements supplémentaires à effectuer.

Deux projets débutent sur ces thèmes, en lien avec des océanographes. Ils ne sont pas sans rapport avec l'étude de sensibilité évoquée précédemment ; en effet il s'agit dans un cas de contrôler des paramètres du schéma de mélange vertical et dans l'autre une partie de la paramétrisation des échanges air/mer. Le premier de ces projets s'intéresse donc au calcul de la profondeur de couche de mélange (MLD), quantité importante comme on l'a vu pour la prévision de la température de surface mais également pour le couplage avec les modèles de biogéochimie car elle influe sur la répartition des nutriments. Des comparaisons entre

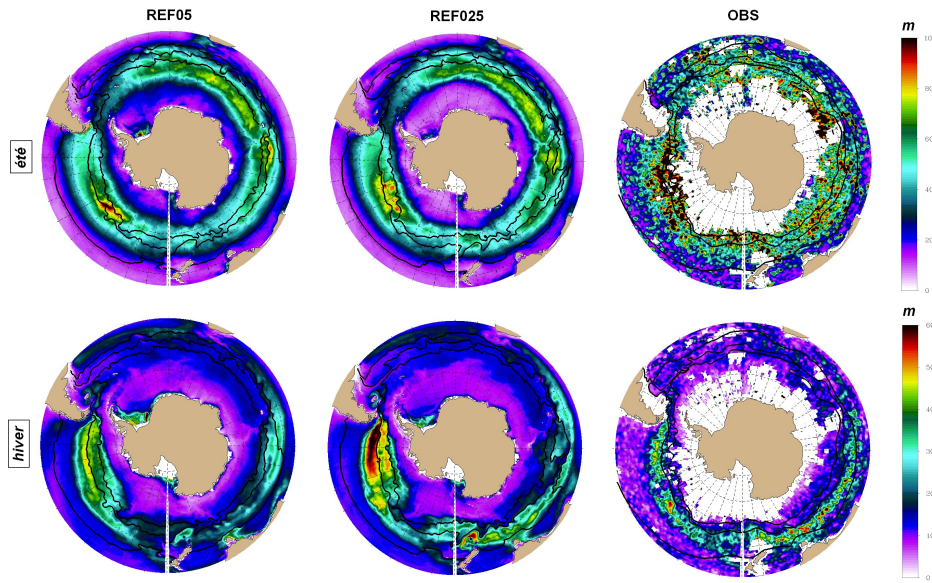


FIGURE 3.3: Profondeurs de couches de mélange (MLD) moyennes (1995-2004) en (haut) été et (bas) hiver, pour des simulations NEMO au $1/2^\circ$, au $1/4^\circ$ et la climatologie de Holte et al. (2010). Figure provenant de la thèse de Carolina Dufour

des climatologies de simulations et de données de MLD (Fig 3.3) mettent en évidence un biais important et dépendant de la saison. Ce biais est vraisemblablement dû à de mauvaises spécifications dans le schéma de mélange vertical. Un contrôle robuste de cette paramétrisation, avec une robustesse aux incertitudes dues aux forçages, devrait permettre d'améliorer grandement le comportement du modèle. L'autre cas d'étude qui m'a été proposé concerne le calcul du "stress" air-ocean où il apparaît qu'utiliser la même formule *bulk* en mode couplé et en mode forcé induit une réduction significative de l'énergie cinétique de surface dans ce dernier cas. Il est donc nécessaire d'appliquer un terme correctif dont on ne connaît pas la variabilité spatiale et temporelle. Ce cadre se prête bien à une étude de sensibilité couplée à un contrôle (robuste ou non) de paramètres et sera une bonne illustration de l'utilité d'effectuer de telles études.

3.4 Analyse de stabilité

Au delà de l'assimilation de données et du contrôle, les modèles tangent et adjoint sont des outils précieux pour diagnostiquer le comportement du modèle direct.

L'océan est un système dynamique chaotique doté d'un attracteur étrange, les trajectoires de ce système sont donc attirées vers une variété de dimension faible. Les perturbations orthogonales à cet attracteur vont donc être naturellement dissipées alors que les perturbations tangentées à l'attracteur seront amplifiées. Pour de nombreuses raisons il peut être intéressant de caractériser ces modes instables. Un moyen d'y parvenir est d'étudier l'évolution de perturbation par le modèle. On a toujours notre modèle :

$$\begin{cases} \frac{dx}{dt} = \mathcal{M}(x), & t \in [0, T], \\ x(0) = x_0, \end{cases} \quad (3.20)$$

On définit classiquement le taux d'accroissement d'une perturbation $\delta\mathbf{x}_0$

$$\rho(\delta\mathbf{x}(0)) = \frac{\|\mathcal{M}_{0 \rightarrow T}(\mathbf{x}(0) + \delta\mathbf{x}(0)) - \mathcal{M}_{0 \rightarrow T}(\mathbf{x}(0))\|}{\|\delta\mathbf{x}(0)\|} \quad (3.21)$$

où $\|\cdot\|$ est une norme donnée.

On peut alors définir la perturbation optimale $\delta\mathbf{x}_1^*(0)$ telle que $\rho(\delta\mathbf{x}_1^*(0)) = \max_{\delta\mathbf{x}(0)} \rho(\delta\mathbf{x}(0))$ puis en déduire une famille de vecteur de croissance maximale

$$\rho(\delta\mathbf{x}_i^*(0)) = \max_{\delta\mathbf{x}(0) \perp \text{Span}(\delta\mathbf{x}_1^*(0), \dots, \delta\mathbf{x}_{i-1}^*(0))} \rho(\delta\mathbf{x}(0)) \quad , i \geq 2 \quad (3.22)$$

Dans le cas où \mathcal{M} est linéaire, ou si on en fait l'approximation par le modèle linéaire tangent \mathbf{M} , on est dans le cadre classique des système dynamiques linéaires et le facteur d'accroissement devient :

$$\begin{aligned} \rho^2(\delta\mathbf{x}(0)) &= \frac{\|\mathbf{M}_{0 \rightarrow T}(\delta\mathbf{x}(0))\|^2}{\|\delta\mathbf{x}(0)\|^2} = \frac{\langle \mathbf{M}_{0 \rightarrow T} \delta\mathbf{x}(0), \mathbf{M}_{0 \rightarrow T} \delta\mathbf{x}(0) \rangle}{\langle \delta\mathbf{x}(0), \delta\mathbf{x}(0) \rangle} \\ &= \frac{\langle \delta\mathbf{x}(0), \mathbf{M}_{0 \rightarrow T}^* \mathbf{M}_{0 \rightarrow T} \delta\mathbf{x}(0) \rangle}{\langle \delta\mathbf{x}(0), \delta\mathbf{x}(0) \rangle} \end{aligned} \quad (3.23)$$

$\mathbf{M}_{0 \rightarrow T}^* \mathbf{M}_{0 \rightarrow T}$ étant une matrice symétrique définie positive, ses valeurs propres sont des réels positifs et ses vecteurs propres sont (ou peuvent être choisis) orthogonaux. Les vecteurs de plus forts croissance sont les premiers vecteurs propres de $\mathbf{M}_{0 \rightarrow T}^* \mathbf{M}_{0 \rightarrow T}$, correspondants au plus grandes valeur propres. On les appelle les vecteurs singuliers directs (FSV).

$$\mathbf{M}_{0 \rightarrow T}^* \mathbf{M}_{0 \rightarrow T} f_i^+ = \mu_i f_i^+ \quad (3.24)$$

$\mathbf{M}_{0 \rightarrow T} f_i^+$ est un vecteur propre de $\mathbf{M}_{0 \rightarrow T} \mathbf{M}_{0 \rightarrow T}^*$.

Les vecteurs singuliers rétrogrades (BSV), notés f_i^- , sont définis par:

$$\mathbf{M}_{0 \rightarrow T} f_i^+ = \sqrt{\mu_i} f_i^-$$

La valeur propres correspondant à f_i^- est également μ_i .

Le calcul des f_i^+ et des f_i^- peut être effectué par des algorithmes des décomposition en valeurs propres tel que l'algorithme de Lanczos. Cela nécessite des multiplications matrice-vecteurs, c'est à dire des intégrations directes du modèle et rétrograde de l'adjoint. Le résultat de ces calculs dépend de la norme utilisée, de la fenêtre temporelle et de l'état initial utilisé si le modèle est non linéaire (on est à nouveau dans une étude locale car on différentie autour d'une trajectoire de référence). Si on étend ce concept à un temps infini on retrouve alors les vecteurs de Llyapunov. Si on considère l'évolution non linéaire, les Llyapunov correspondent aux breeding vectors (Toth and Kalnay, 1997) et les vecteurs singuliers aux vecteurs singuliers non linéaires (Mu et al., 2000). Le calcul de ces derniers se fait par minimisations successives et peut donner des perturbations optimales non physiques ou infinitésimales. Pour pallier ces déficiences on peut effectuer ces minimisations sous contrainte, on obtient alors les Constrained Non-linear Optimal Perturbations (Mu et al., 2003; Rivière et al., 2009). Cette liste de vecteurs caractéristiques n'est pas exhaustive, on peut par exemple en trouver d'autres dans les travaux de Moore et al. (2004).

Ces vecteurs, par l'information qu'ils contiennent sur le comportement du système ont de nombreuses applications, outre les études du fonctionnement même de ces système, on peut citer la prévision d'ensemble, les études de sensibilité (Rivière et al. (2009) pour une application récente), la réduction d'ordre en

assimilation de données (Durbiano, 2001; Blayo et al., 2003), pour améliorer le réseau d'observation (Qin and Mu, 2011) ou permettre de mieux choisir les observation ciblées (Mu et al., 2009). Hormis la génération d'ensemble, cependant, le potentiel de ces vecteurs a encore été assez peu exploité. Nous sommes actuellement en train de mettre en place de tels outils pour la plateforme NEMO dans le cadre de la thèse de P.-A. Bouttier dans le but de fournir des diagnostics novateurs pour la modélisation et l'assimilation de données.

3.5 Conclusion

On a vu dans ce chapitre un panorama des applications des outils variationnels qui fournissent un cadre unifié pour l'analyse de sensibilité et de stabilité, la propagation d'incertitude, l'assimilation de données, et le contrôle de paramètres, le cœur de ces applications étant le calcul de dérivée (sensibilités) locales. On a également illustré le fait que ces outils permettaient un calcul exact de ces sensibilités, mais que cela ne permettait qu'une étude locale ou dans le cas linéaire. L'extension aux aspects globaux et/ou non-linéaires est possible mais coûteuse et souvent au prix d'approximations et définir des méthodes performantes est un sujet d'actualité.

Les approches proposées ici ont un fort potentiel dans l'amélioration des systèmes de prévision, tant pour la partie modélisation que la partie assimilation. On peut d'ailleurs concevoir certaines des méthodes d'assimilation de données variationnelles d'ensemble évoquées au chapitre 1 comme une approche robuste particulière de l'assimilation variationnelle ; cette analogie pourrait nous permettre d'élaborer d'autres extensions de méthodes variationnelle d'assimilation de données pour prendre en compte les incertitudes, mais ceci est une autre histoire ...

Bibliographie

- Adrian, R. J., 1991: Particle-imaging techniques for experimental fluid mechanics. *Annual review of fluid mechanics*, **23** (1), 261–304.
- Amodei, L. and M. N. Benbourhim, 1991: A vector spline approximation. *J. Approx. Theory*, **67** (1), 1–29.
- Anderson, D. L. T., et al., 2003: Comparison of the ECMWF seasonal forecast systems 1 and 2, including the relative performance for the 1997/8 El Nino. . *ECMWF Technical Memorandum*, **404**, 1–93.
- Andersson, E. and H. Järvinen, 1999: Variational quality control. *Q.J.R. Meteorol. Soc.*, **125** (554), 1–26.
- Artale, V., G. Boffetta, A. Celani, M. Cencini, and A. Vulpiani, 1997: Dispersion of passive tracers in closed basins: Beyond the diffusion coefficient. *Phys. Fluids*, **9** (11), 3162–3171.
- Aubert, G. and P. KornProbst, 2006: *Mathematical Problems in Image Processing*. Springer, New York.
- Aurell, E., G. Boffetta, A. Crisanti, G. Paladin, and A. Vulpiani, 1997: Predictability in the large: an extension of the concept of Lyapunov exponent. *Journal of Physics A: Mathematical and General*, **30** (1), 1–26.
- Auroux, D. and J. BLUM, 2008: A nudging-based data assimilation method: the Back and Forth Nudging (BFN) algorithm. *Nonlinear Processes in Geophysics*, **15**, 305–319.
- Auroux, D. and J. Fehrenbach, 2010: Identification of velocity fields for geophysical fluids from a sequence of images. *Exp Fluids*, **50** (2), 313–328.
- Balmaseda, M. A., D. L. T. Anderson, and A. Vidard, 2007a: Impact of Argo on analyses of the global ocean. *Geophys. Res. Lett.*, **34** (16).
- Balmaseda, M. A., D. P. Dee, A. Vidard, and D. L. T. Anderson, 2007b: A multivariate treatment of bias for sequential data assimilation: Application to the tropical oceans. *Q.J.R. Meteorol. Soc.*, **133** (622), 167–179.
- Balmaseda, M. A., G. C. Smith, K. Haines, D. L. T. Anderson, T. N. Palmer, and A. Vidard, 2007c: Historical reconstruction of the Atlantic Meridional Overturning Circulation from the ECMWF operational ocean reanalysis. *Geophys. Res. Lett.*, **34** (23).
- Balmaseda, M. A., A. Vidard, and D. L. T. Anderson, 2007d: Climate variability from the new System 3 ocean analysis. *ECMWF newsletter*, **113** (Autumn issue).
- Balmaseda, M. A., A. Vidard, and D. L. T. Anderson, 2007e: The ECMWF System 3 ocean analysis. *ECMWF Technical Memorandum*, **508**, 1–47.

-
- Balmaseda, M. A., A. Vidard, and D. L. T. Anderson, 2008: The ECMWF Ocean Analysis System: ORA-S3. *Mon. Wea. Rev.*, **136** (8), 3018–3034.
- Bengtsson, T., C. Snyder, and D. Nychka, 2003: Toward a nonlinear ensemble filter for high-dimensional systems. *J. Geophys. Res.*, **108** (D24), 8775–.
- Beron-Vera, F. J., 2010: Mixing by low- and high-resolution surface geostrophic currents. *J. Geophys. Res.*, **115** (C10027), 15pp.
- Beron-Vera, F. J. and M. J. Olascoaga, 2009: An Assessment of the Importance of Chaotic Stirring and Turbulent Mixing on the West Florida Shelf. *J. Phys. Oceanogr.*, **39** (7), 1743–1755.
- Beron-Vera, F. J., M. J. Olascoaga, and G. J. Goni, 2010: Surface Ocean Mixing Inferred from Different Multisatellite Altimetry Measurements. *J. Phys. Oceanogr.*, **40** (11), 2466–2480.
- Birol, F., J. M. Brankart, J.-M. Lemoine, P. Brasseur, and J. Verron, 2005: Assimilation of satellite altimetry referenced to the new GRACE geoid estimate. *Geophys. Res. Lett.*, **32**.
- Blayo, E., S. Durbiano, A. Vidard, and F.-X. Le Dimet, 2003: Reduced order strategies for variational data assimilation in oceanic models. *Data Assimilation for Geophysical Flows.*, B. Sportisse and F.-X. Le Dimet, Eds., Springer-Verlag, 18pp.
- Boilley, A. and J.-F. Mahfouf, 2012: Assimilation of low-level wind in a high-resolution mesoscale model using the back and forth nudging algorithm. *Tellus A*, **64** (0).
- Bormann, N. and P. Bauer, 2010: Estimates of spatial and interchannel observation-error characteristics for current sounder radiances for numerical weather prediction. I: Methods and application to ATOVS data. *Q.J.R. Meteorol. Soc.*, **136** (649), 1036–1050.
- Bormann, N., A. Collard, and P. Bauer, 2010: Estimates of spatial and interchannel observation-error characteristics for current sounder radiances for numerical weather prediction. II: Application to AIRS and IASI data. *Q.J.R. Meteorol. Soc.*, **136** (649), 1051–1063.
- Bormann, N., S. Saarinen, and G. Kelly, 2003: The spatial structure of observation errors in atmospheric motion vectors from geostationary satellite data. *Mon. Wea. Rev.*, **131** (4), 706–718.
- Brankart, J.-M., E. Cosme, C.-E. Testut, P. Brasseur, and J. Verron, 2010: Efficient Adaptive Error Parameterizations for Square Root or Ensemble Kalman Filters: Application to the Control of Ocean Mesoscale Signals. *Mon. Wea. Rev.*, **138** (3), 932–950.
- Brankart, J.-M., E. Cosme, C.-E. Testut, P. Brasseur, and J. Verron, 2011: Efficient Local Error Parameterizations for Square Root or Ensemble Kalman Filters: Application to a Basin-Scale Ocean Turbulent Flow. *Mon. Wea. Rev.*, **139** (2), 474–493.
- Brasseur, P., 2006: Ocean data assimilation using sequential methods based on the Kalman Filter. *Ocean weather forecasting: An integrated view of oceanography.*, J. Verron and É. Chassignet, Eds., Springer, Dordrecht, 271–316.
- Brasseur, P. and J. Verron, 2006: The SEEK filter method for data assimilation in oceanography: a synthesis. *Ocean Dynamics*, **56** (5-6), 650–661.

-
- Brasseur, P., et al., 2005: Data assimilation for marine monitoring and prediction: The MERCATOR operational assimilation systems and the MERSEA developments. *Q.J.R. Meteorol. Soc.*, **131 (613)**, 3561–3582.
- Buehner, M., P. L. Houtekamer, C. Charette, H. L. Mitchell, and B. He, 2010a: Intercomparison of Variational Data Assimilation and the Ensemble Kalman Filter for Global Deterministic NWP. Part I: Description and Single-Observation Experiments. *Mon. Wea. Rev.*, **138 (5)**, 1550–1566.
- Buehner, M., P. L. Houtekamer, C. Charette, H. L. Mitchell, and B. He, 2010b: Intercomparison of Variational Data Assimilation and the Ensemble Kalman Filter for Global Deterministic NWP. Part II: One-Month Experiments with Real Observations. *Mon. Wea. Rev.*, **138 (5)**, 1567–1586.
- Cacuci, D. G., 2003: *Sensitivity and Uncertainty Analysis: Theory*. CRC Press.
- Candès, E. and D. L. Donoho, 2003: New tight frames of curvelets and optimal representations of objects with piecewiseC2 singularities. *Comm. Pure Appl. Math.*, **57 (2)**, 219–266.
- Castruccio, F., J. Verron, L. Gourdeau, J. M. Brankart, and P. Brasseur, 2006: On the role of the GRACE mission in the joint assimilation of altimetric and TAO data in a tropical Pacific Ocean model. *Geophys. Res. Lett.*, **33**.
- Chan, T. F. and J. Shen, 2005: *Image processing and analysis. variational, PDE, wavelet, and stochastic methods*, SIAM, Philadelphia.
- Chastaing, G., F. Gamboa, and C. Prieur, 2012: Generalized Hoeffding-Sobol Decomposition for Dependent Variables- Application to Sensitivity Analysis. *Electronic Journal of Statistics*, 1–34.
- Chauvin, C., M. Nodet, A. Vidard, and P.-A. Bouttier, 2011: Assimilation of Lagrangian Data in an operational framework. Tech. rep., INRIA.
- Chu, P. C., L. Ivanov, and T. Margolina, 2007: On non-linear sensitivity of marine biological models to parameter variations. *Ecological modelling*, **206 (3-4)**, 369–382.
- Church, J. A. and N. J. White, 2006: A 20th century acceleration in global sea-level rise. *Geophys. Res. Lett.*, **33 (1)**.
- Corpetti, T., P. Héas, É. Memin, and N. Papadakis, 2009: Pressure image assimilation for atmospheric motion estimation. *Tellus A*, **61 (1)**, 160–178.
- Courtier, P., J.-N. Thépaut, and A. Hollingsworth, 1994: A strategy for operational implementation of 4D-Var, using an incremental approach. *QJR Meteorol Soc*, **120 (519)**, 1–21.
- Davey, M., et al., 2006: Multi-model multi-method multi-decadal ocean analyses from the ENACT project. *Clivar exchange*, **11 (3)**, 22–25.
- Dee, D. P. and S. Uppala, 2009: Variational bias correction of satellite radiance data in the ERA-Interim reanalysis. *Q.J.R. Meteorol. Soc.*, **135 (644)**, 1–12.
- Derber, J. and F. Bouttier, 1999: A reformulation of the background error covariance in the ECMWF global data assimilation system. *Tellus A*, **51 (2)**, 195–221.

-
- d'Ovidio, F., V. Fernández, E. Hernández-García, and C. López, 2004: Mixing structures in the Mediterranean Sea from finite-size Lyapunov exponents. *Geophys. Res. Lett.*, **31** (L17203), 4pp.
- d'Ovidio, F., J. Isern-Fontanet, C. López, E. Hernández-García, and E. García-Ladona, 2009a: Comparison between Eulerian diagnostics and finite-size Lyapunov exponents computed from altimetry in the Algerian basin. *Deep Sea Research Part I: Oceanographic Research Papers*, **56** (1), 15–31.
- d'Ovidio, F., V. Taillandier, I. Taupier-Letage, and L. Mortier, 2009b: Lagrangian validation of the Mediterranean mean dynamic topography by extraction of tracer frontal structures. *Mercator Ocean Quarterly Newsletter*, **32**, 24–32.
- Durbiano, S., 2001: Vecteurs caractéristiques de modèles océaniques pour la réduction d'ordre en assimilation de données. Ph.D. thesis, Université Joseph Fourier.
- Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.*, **99** (C5), 10 143–10 162.
- Evensen, G., 2003: The Ensemble Kalman Filter: theoretical formulation and practical implementation. *Ocean Dynamics*, **53** (4), 343–367.
- Ferron, B., 2011: A 4D-variational approach applied to an eddy-permitting North Atlantic configuration: synthetic and real data assimilation of altimeter observations. *Ocean Modelling*, **39**, 370–385.
- Fertig, E. J., J. Harlim, and B. R. Hunt, 2007: A comparative study of 4D-VAR and a 4D Ensemble Kalman Filter: perfect model simulations with Lorenz-96. *Tellus A*, **59** (1), 96–100.
- Fisher, M. and H. Auvinen, 2012: Long Window 4D-Var. *Proceedings of the ECMWF Seminar Series on Data assimilation for atmosphere and ocean, 6-9 September 2011*, 189–202.
- Fu, L.-L. and A. Cazenave, 2001: *Satellite altimetry and earth sciences*, A handbook of techniques and applications, Vol. 69. ACADEMIC PRESS.
- Gao, J. and M. Xue, 2008: An Efficient Dual-Resolution Approach for Ensemble Data Assimilation and Tests with Simulated Doppler Radar Data. *Mon. Wea. Rev.*, **136** (3), 945–963.
- Gaultier, L., J. Verron, J.-M. Brankart, O. Titaud, and P. Brasseur, 2012: On the inversion of submesoscale tracer fields to estimate the surface ocean circulation. *Journal of Marine Systems*, in press.
- Gejadze, I. Y., G. J. M. Copeland, F.-X. Le Dimet, and V. Shutyaev, 2011: Computation of the analysis error covariance in variational data assimilation problems with nonlinear dynamics. *Journal of Computational Physics*, **230** (22), 7923–7943.
- Geman, S. and D. Geman, 1984: Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images. *IEEE Trans. Pattern Anal. Mach. Intell.*, **6**, 721–741.
- Greatbatch, R. J., 1994: A note on the representation of steric sea level in models that conserve volume rather than mass. *J. Geophys. Res.*, **99** (C6), 12 767–12 771.
- Griffith, A. K. and N. K. Nichols, 2000: Adjoint methods in data assimilation for estimating model error. *Flow, Turbulence and Combustion*, **65** (3/4), 469–488.

-
- Haben, S. A., A. S. Lawless, and N. K. Nichols, 2011: Conditioning and preconditioning of the variational data assimilation problem. *Computers & Fluids*, **46** (1), 252–256.
- Haines, K., J. D. Blower, J. P. Drecourt, A. Vidard, I. Astin, and X. Zhou, 2006: Salinity assimilation using S (T): Covariance relationships. *Mon. Wea. Rev.*, **134** (3), 759–771.
- Haller, G., 2000: Finding finite-time invariant manifolds in two-dimensional velocity fields. *Chaos*, **10** (1), 99–108.
- Haller, G., 2001: Lagrangian structures and the rate of strain in a partition of two-dimensional turbulence. *Phys. Fluids*, **13** (11), 3365–3385.
- Haller, G., 2002: Lagrangian coherent structures from approximate velocity data. *Phys. Fluids*, **14** (6), 1851–1861.
- Haller, G., 2011: A variational theory of hyperbolic Lagrangian Coherent Structures. *Physica D*, **240** (7), 574–598.
- Haller, G. and G. Yuan, 2000: Lagrangian coherent structures and mixing in two-dimensional turbulence. *Physica D: Nonlinear Phenomena*, **147** (3-4), 352–370.
- Hamill, T. M. and C. Snyder, 2010: A Hybrid Ensemble Kalman Filter–3D Variational Analysis Scheme. *Mon. Wea. Rev.*, **128** (8), 2905–2919.
- Homma, T. and A. Saltelli, 1996: Importance measures in global sensitivity analysis of nonlinear models. *Reliability Engineering & System Safety*, **52** (1), 1–17.
- Horn, B. K. P. and B. G. Schunck, 1981: Determining optical flow. *Artificial Intelligence*, **17** (1-3), 185–203.
- Hoteit, I., X. Luo, and D. T. Pham, 2011: Particle Kalman Filtering: A Nonlinear Bayesian Framework for Ensemble Kalman Filters. *Mon. Wea. Rev.*, 110823130907004.
- Hoteit, I., D. Pham, and G. Triantafyllou, 2008a: Particle Kalman filtering for data assimilation in meteorology and oceanography. *Proceedings of the 3rd International Conference on Reanalysis (WCRP '08)*, Tokyo, Japan, 1–6.
- Hoteit, I., D.-T. Pham, G. Triantafyllou, and G. Korres, 2008b: A New Approximate Solution of the Optimal Nonlinear Filter for Data Assimilation in Meteorology and Oceanography. *Mon. Wea. Rev.*, **136** (1), 317–334.
- Hunt, B. R., et al., 2004: Four-dimensional ensemble Kalman filtering. *Tellus A*, **56** (4), 273–277.
- Ingleby, B. and M. Huddleston, 2007: Quality control of ocean temperature and salinity profiles — Historical and real-time data. *Journal of Marine Systems*, **65** (1-4), 158–175.
- Ingleby, B. and A. C. Lorenc, 1993: Bayesian quality control using multivariate normal distributions. *Q.J.R. Meteorol. Soc.*, **119** (513), 1195–1225.
- Jazwinski, A. H., 1970: *Stochastic processes and filtering theory*. Academic Press, New York.
- Kalman, R., 1960: A new approach to linear filtering and prediction problems. *Journal of basic Engineering*.

-
- Kalnay, E., H. LI, T. Miyoshi, S.-C. YANG, and J. Ballabrera-Poy, 2007: 4-D-Var or ensemble Kalman filter? *Tellus A*, **59** (5), 758–773.
- Korotaev, G., E. Huot, F. Le Dimet, I. Herlin, S. Stanichny, D. Solovyeu, and L. Wu, 2008: Retrieving ocean surface current by 4-D variational assimilation of sea surface temperature images. *Remote Sensing of Environment*, **112** (4), 1464–1475.
- Kraus, T., P. Kuhl, L. Wirsching, H. G. Bock, and M. Diehl, 2006: A moving horizon state estimation algorithm applied to the tennessee eastman benchmark process. *IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems*, Heidelberg, Germany, 377–382.
- Krysta, M., E. Blayo, E. Cosme, and J. Verron, 2011: A consistent hybrid variational-smoothing data assimilation method: Application to a simple shallow-water model of the turbulent mid-latitude ocean. *Mon. Wea. Rev.*, 110506143106037.
- Krysta, M., E. Blayo, E. Cosme, J. Verron, and A. Vidard, 2008: A 4D-Var SEEK smoother hybrid. Towards applications in oceanography. *WWRP/THORPEX Workshop on 4D-Var and Ensemble Kalman filter inter-comparisons.*, Buenos-Aires, Argentina.
- Kucherenko, S., M. Rodriguez-Fernandez, C. Pantelides, and N. Shah, 2009: Monte Carlo evaluation of derivative-based global sensitivity measures. *Reliability Engineering & System Safety*, **94** (7), 1135–1148.
- Lagarde, T., 2000: Nouvelle approche des méthodes d’assimilation de données: les algorithmes de point selle. Ph.D. thesis, Université Paul Sabatier, Toulouse.
- Lapeyre, G., 2002: Characterization of finite-time Lyapunov exponents and vectors in two-dimensional turbulence. *Chaos*, **12** (3), 688–11.
- Lauvernet, C., J. M. Brankart, F. Castruccio, G. Broquet, P. Brasseur, and J. Verron, 2009: A truncated Gaussian filter for data assimilation with inequality constraints: Application to the hydrostatic stability condition in ocean models. *Ocean Modelling*, **27** (1-2), 1–17.
- Le Dimet, F.-X. and O. Talagrand, 1986: Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects. *Tellus A*, **38A** (2), 97–110.
- Lea, D. J., J.-P. Drécourt, K. Haines, and M. J. Martin, 2008: Ocean altimeter assimilation with observational- and model-bias correction. *Q.J.R. Meteorol. Soc.*, **134** (636), 1–14.
- Lehahn, Y., F. d’Ovidio, M. Lévy, and E. Heifetz, 2007: Stirring of the northeast Atlantic spring bloom: A Lagrangian analysis based on multisatellite data. *J. Geophys. Res.*, **112** (C8), 15.
- Lemieux, B. and A. Vidard, 2011: Weak constraint 4D-Var in NEMOVAR, a multiple snapshot approach. *VODA report*, **D3.2.1**.
- Lévy, M., D. Iovino, S. Masson, G. Madec, P. Klein, A. M. Tréguier, and K. Takahashi, 2009: Remote impacts of Sub-Mesoscale Dynamics on new production. *Mercator Ocean Quarterly Newsletter*, **13**, 13–17.
- Liu, M., 1995: *Variational assimilation of acoustic tomography*. Technical report (Florida State University. Center for Ocean-Atmospheric Prediction Studies), Florida State University.

-
- Liu, Z. Q. and F. Rabier, 2006: The potential of high-density observations for numerical weather prediction: A study with simulated observations. *Q.J.R. Meteorol. Soc.*, **129** (594), 3013–3035.
- Lorenc, A. C., 1981: A global three-dimensional multivariate statistical interpolation scheme. *Mon. Wea. Rev.*, **109** (4), 701–721.
- Luong, B., J. Blum, and J. Verron, 1998: A variational method for the resolution of a data assimilation problem in oceanography. *Inverse Problems*, 1–20.
- Ma, J., O. Titaut, A. Vidard, and F.-X. Le Dimet, 2008: Spatio-temporal structure extraction and denoising of geophysical fluid image sequences using 3D curvelet transforms. *INRIA Research report*, **6683**, 1–30.
- Mahalanobis, P. C., 1936: On the generalised distance in statistics. *Proceedings of the National Institute of Sciences of India*, **2** (1), 49–55.
- Mallat, S. G., 1998: *A wavelet tour of signal processing*. Morgan Kaufmann.
- Mathur, M., G. Haller, T. Peacock, J. E. Ruppert-Felsot, and H. L. Swinney, 2007: Uncovering the Lagrangian Skeleton of Turbulence. *Phys. Rev. Lett.*, **98** (14), 144 502.
- Maximenko, N., P. Niiler, L. Centurioni, M.-H. Rio, O. Melnichenko, D. Chambers, V. Zlotnicki, and B. Galperin, 2009: Mean Dynamic Topography of the Ocean Derived from Satellite and Drifting Buoy Data Using Three Different Techniques*. *J. Atmos. Oceanic Technol.*, **26** (9), 1910–1919.
- McPhaden, M. J., et al., 1998: The Tropical Ocean-Global Atmosphere observing system: A decade of progress. *J. Geophys. Res.*, **103** (C7), 14 169–14 240.
- Michel, Y., 2011: Displacing Potential Vorticity Structures by the Assimilation of Pseudo-Observations. *Mon. Wea. Rev.*, **139** (2), 549–565.
- Michel, Y. and F. Bouttier, 2006: Automated tracking of dry intrusions on satellite water vapour imagery and model output. *Q.J.R. Meteorol. Soc.*, **132** (620), 2257–2276.
- Mirouze, I. and A. T. Weaver, 2010: Representation of correlation functions in variational assimilation using an implicit diffusion operator. *Q.J.R. Meteorol. Soc.*, **136** (651), 1421–1443.
- Mogensen, K. S., M. A. Balmaseda, M. Martin, and A. Vidard, 2009: NEMOVAR: A variational data assimilation system for the NEMO ocean model. *ECMWF newsletter*, **120**, 17–21.
- Molteni, F., R. Buizza, T. N. Palmer, and T. Petroliaigis, 1996: The ECMWF ensemble prediction system: methodology and validation, QJ Roy. *Q.J.R. Meteorol. Soc.*, **122** (529), 73–119.
- Moore, A. M., H. G. Arango, E. Di Lorenzo, B. Cornuelle, A. Miller, and D. Neilson, 2004: A comprehensive ocean prediction and analysis system based on the tangent linear and adjoint of a regional ocean model. *Ocean Modelling*, **7** (1-2), 227–258.
- Morris, M., 1991: Factorial sampling plans for preliminary computational experiments. *Technometrics*, **33** (2), 161–174.
- Mu, M., W. Duan, and B. Wang, 2003: Conditional nonlinear optimal perturbation and its applications. *Nonlinear Processes in Geophysics*, **10** (6), 493–501.

-
- Mu, M., G. Huan, W. Jiafeng, and L. Yong, 2000: The impact of nonlinear stability and instability on the validity of the tangent linear model. *Adv. Atmos. Sci.*, **17** (3), 375–390.
- Mu, M., F. Zhou, and H. Wang, 2009: A Method for Identifying the Sensitive Areas in Targeted Observations for Tropical Cyclone Prediction: Conditional Nonlinear Optimal Perturbation. *Mon. Wea. Rev.*, **137** (5), 1623–1639.
- Nerger, L., W. Hiller, and J. Schröter, 2005: A comparison of error subspace Kalman filters. *Tellus A*, **57** (5), 715–735.
- Neveu, É., 2011: Application des méthodes multigrilles à l’assimilation variationnelle de données en géophysique. Ph.D. thesis, Université de Grenoble.
- Nieman, S., P. MENzel, C. Hayden, D. Gray, S. Wanzong, C. Velden, and J. Daniels, 1997: WP Menzel, CM Hayden, D. Gray, ST Wanzong, CS Velden, and J. Daniels, 1997: Fully automated cloud-drift winds in NESDIS operations. *Bull. Amer. Meteor. Soc.*, **78**, 1121–1133.
- Nocedal, J. and S. J. Wright, 1999: Numerical optimization. Springer Verlag.
- Nodet, M., 2006: Variational assimilation of Lagrangian data in oceanography. *Inverse Problems*, **22**, 245.
- Nodet, M., 2009: Optimal control of the Primitive Equations of the ocean with Lagrangian observations. *ESAIM: COCV*, **16** (2), 400–419.
- Olascoaga, M. J., F. J. Beron-Vera, L. E. Brand, and H. K. c. c. ak, 2008: Tracing the early development of harmful algal blooms on the West Florida Shelf with the aid of Lagrangian coherent structures. *J. Geophys. Res.*, **113** (C12014), 10 pp.
- Olascoaga, M. J., I. I. Rypina, M. G. Brown, F. J. Beron-Vera, H. K. c. c. ak, L. E. Brand, G. R. Halliwell, and L. K. Shay, 2006: Persistent transport barrier on the West Florida Shelf. *Geophys. Res. Lett.*, **33** (L22603), 5pp.
- Ott, E., 1993: *Chaos in Dynamical Systems*. Chaos in Dynamical Systems, Cambridge University Press, New York, USA.
- Ott, E., et al., 2004: A Local Ensemble Kalman Filter for Atmospheric Data Assimilation. *Tellus A*, **56** (5), 415–428.
- Ottino, J. M., 1989: *The Kinematics of Mixing: Stretching, Chaos, and Transport*. Cambridge University Press.
- Papadakis, N. and É. Memin, 2008: Variational Assimilation of Fluid Motion from Image Sequence. *SIAM J. on Imaging Sciences*, **1** (4), 343–363.
- Pham, D. T., 1996: A singular evolutive interpolated Kalman filter for data assimilation in oceanography. *Rapport technique LMC/IMAG*, 1–22.
- Pham, D. T., 2001: Stochastic Methods for Sequential Data Assimilation in Strongly Nonlinear Systems. *Mon. Wea. Rev.*, **129** (5), 1194–1207.
- Pham, D. T., J. Verron, and M.-C. Roubaud, 1998: A singular evolutive extended Kalman filter for data assimilation in oceanography. *Journal of Marine Systems*, **16** (3-4), 323–340.

-
- Pires, C., R. Vautard, and T. Olivier, 1996: On extending the limits of variational assimilation in nonlinear chaotic systems. *Tellus A*, **48** (1), 96–121.
- Qin, X. and M. Mu, 2011: Influence of conditional nonlinear optimal perturbations sensitivity on typhoon track forecasts. *Q.J.R. Meteorol. Soc.*, n/a–n/a.
- Rabier, F., 2005: Overview of global data assimilation developments in numerical weather-prediction centres. *Q.J.R. Meteorol. Soc.*, **131** (613), 3215–3233.
- Raynaud, L., L. Berre, and G. Desroziers, 2011: An extended specification of flow-dependent background error variances in the Météo-France global 4D-Var system. *Q.J.R. Meteorol. Soc.*, **137** (656), 607–619.
- Richman, J. G., R. N. Miller, and Y. Spitz, 2005: Error estimates for assimilation of satellite sea surface temperature data in ocean climate models. *Geophys. Res. Lett.*, **32** (L18608).
- Rio, M. H. and F. Hernandez, 2004: A mean dynamic topography computed over the world ocean from altimetry, in situ measurements, and a geoid model. *J. Geophys. Res.*, **109** (C12), C12 032–.
- Rivière, O., G. Lapeyre, and O. Talagrand, 2009: A novel technique for nonlinear sensitivity analysis: application to moist predictability. *Q.J.R. Meteorol. Soc.*, **135** (643), 1520–1537.
- Saltelli, A., K. Chan, and E. Scott, 2000: Sensitivity analysis. Wiley series in probability and statistics, Chichester.
- Schmetz, J., K. Holmlund, J. Hoffman, B. Strauss, B. Mason, V. Gaertner, A. Koch, and B. van DE, 1993: Operational cloud-motion winds from Meteosat infrared images. *Journal of applied meteorology*, **32** (7), 1206–1225.
- Shadden, S. C., F. Lekien, and J. E. Marsden, 2005: Definition and properties of Lagrangian coherent structures from finite-time Lyapunov exponents in two-dimensional aperiodic flows. *Physica D: Nonlinear Phenomena*, **212** (3-4), 271–304.
- Shadden, S. C., F. Lekien, J. D. Paduan, F. P. Chavez, and J. E. Marsden, 2009: The correlation between surface drifters and coherent structures based on high-frequency radar data in Monterey Bay. *Deep Sea Research Part II: Topical Studies in Oceanography*, **56** (3-5), 161–172.
- Simon, E. and L. Bertino, 2009: Application of the Gaussian anamorphosis to assimilation in a 3-D coupled physical-ecosystem model of the North Atlantic with the EnKF: a twin experiment. *Ocean Sci.*, **5** (4), 495–510.
- Smith, N. R., J. E. Blomley, and G. Meyers, 1991: A univariate statistical interpolation scheme for subsurface thermal analyses in the tropical oceans. *Progress In Oceanography*, **28** (3), 219–256.
- Sobol, I. M., 1990: Sensitivity estimates for nonlinear mathematical models. *Matem. Mod. (in russian)*, **2** (1).
- Sobol, I. M. and S. Kucherenko, 2009: Derivative based global sensitivity measures and their link with global sensitivity indices. *Mathematics and Computers in Simulation*, **79** (10), 3009–3017.
- Souopgui, I., 2010: Assimilation d’images pour les fluides géophysiques. Ph.D. thesis, Université de Grenoble et université de Yaoundé I.

-
- Stepanov, V. N., K. Haines, and G. C. Smith, 2012: Assimilation of RAPID array observations into an ocean model. *Q.J.R. Meteorol. Soc.*, n/a–n/a.
- Stewart, L., S. Dance, and N. Nichols, 2008: Correlated observation errors in data assimilation. *Int. J. Numer. Meth. Fluids*, **56** (8), 1521–1527.
- Storto, A., S. Dobricic, S. Masina, and P. Di Pietro, 2011: Assimilating Along-Track Altimetric Observations through Local Hydrostatic Adjustment in a Global Ocean Variational Assimilation System. *Mon. Wea. Rev.*, **139** (3), 738–754.
- Tapley, B. D., D. P. Chambers, S. Bettadpur, and J. Ries, 2003: Large scale ocean circulation from the GRACE GGM01 Geoid. *Geophys. Res. Lett.*, **30** (22).
- Tikhonov, A., 1963: Regularization of incorrectly posed problems. *Soviet Math. Dokl.*, **4**, 1624–1627.
- Titaut, O., J.-M. Brankart, and J. Verron, 2011: On the use of Finite-Time Lyapunov Exponents and Vectors for direct assimilation of tracer images into ocean models. *Tellus A*, **63** (5), 1038–1051.
- Titaut, O., A. Vidard, I. Souopgui, and F.-X. Le Dimet, 2010: Assimilation of image sequences in numerical models. *Tellus A*, **62** (1), 30–47.
- Toth, Z. and E. Kalnay, 1997: Ensemble forecasting at NCEP and the breeding method. *Mon. Wea. Rev.*, **125** (12), 3297–3319.
- Toye Mahamadou Kele, H., 2009: Interfacing image processing libraries with geophysical fluid forecasting numerical systems. Tech. rep., ENSIMAG.
- Trémolet, Y., 2007a: Accounting for an imperfect model in 4D-Var. *Q.J.R. Meteorol. Soc.*, **132** (621), 2483–2504.
- Trémolet, Y., 2007b: Incremental 4D-Var convergence study. *Tellus A*, **59** (5), 706–718.
- Trémolet, Y., 2007c: Model-error estimation in 4D-Var. *Q.J.R. Meteorol. Soc.*, **133** (626), 1267–1280.
- Van Leeuwen, P. J. and M. Ades, 2012: Efficient fully nonlinear data assimilation for geophysical fluid dynamics. *Computers & Geosciences*, 1–36.
- Veersé, F., D. T. Pham, and J. Verron, 2000: 4D-Var/SEEK: a consistent hybrid variational-smoothing data assimilation method. *INRIA Research report*, **3902**, 1–24.
- Veersé, F. and J.-N. Thépaut, 1998: Multiple-truncation incremental approach for four-dimensional variational data assimilation. *Q.J.R. Meteorol. Soc.*, **124** (550), 1889–1908.
- Verlaan, M. and A. W. Heemink, 1997: Tidal flow forecasting using reduced rank square root filters. *Stochastic Hydrology and Hydraulics*, **11** (5), 349–368.
- Verlaan, M. and A. W. Heemink, 2001: Nonlinearity in Data Assimilation Applications: A Practical Method for Analysis. *Mon. Wea. Rev.*, **129** (6), 1578–1589.
- Vernieres, G., C. K. R. T. Jones, and K. Ide, 2011: Capturing eddy shedding in the Gulf of Mexico from Lagrangian observations. *Physica D*, **240** (2), 166–179.

-
- Vidard, A., 2001: Vers une prise en compte des erreurs-modèle en assimilation de données 4D-variationnelle. Application à un modèle réaliste d'océan. Ph.D. thesis, Université Joseph Fourier.
- Vidard, A., 2010: Simplification operator and its inverse for multi-incremental assimilation. *VODA report*, **D 3.1.1**, 1–15.
- Vidard, A., D. L. T. Anderson, and M. A. Balmaseda, 2007: Impact of Ocean Observation Systems on Ocean Analysis and Seasonal Forecasts. *Mon. Wea. Rev.*, **135 (2)**, 409–429.
- Vidard, A., M. A. Balmaseda, and D. L. T. Anderson, 2009a: Assimilation of Altimeter Data in the ECMWF Ocean Analysis System 3. *Mon. Wea. Rev.*, **137 (4)**, 1393–1408.
- Vidard, A., E. Blayo, F.-X. Le Dimet, and A. Piacentini, 2000: 4D Variational Data Analysis with Imperfect Model. *Flow, Turbulence and Combustion*, **65 (3-4)**, 489–504.
- Vidard, A., F.-X. Le Dimet, and A. Piacentini, 2003: Determination of optimal nudging coefficients. *Tellus A*, **55 (1)**, 1–15.
- Vidard, A., A. Piacentini, and F.-X. Le Dimet, 2004: Variational data analysis with control of the forecast bias. *Tellus A*, **56 (3)**, 1–12.
- Vidard, A., F. Vigilant, C. Deltel, and R. Benshila, 2009b: NEMOTAM Short and long term development strategies. *VODA report*, **D1.1**, 1–14.
- Vidard, A., F. Vigilant, C. Deltel, and R. Benshila, 2010: NEMO Tangent & Adjoint Models (NemoTam) Reference Manual & User's Guide. *VODA report*, **D 1.2**, 1–38.
- Vidard, A., F. Vigilant, C. Deltel, and R. Benshila, 2012: NEMO Tangent & Adjoint Models Reference Manual & User's Guide. *VODA report*, **D1.3**, 1–36.
- Vigan, X., C. Provost, R. Bleck, and P. Courtier, 2000: Sea surface velocities from sea surface temperature image sequences 1. Method and validation using primitive equation model output. *J. Geophys. Res.*, **105 (C8)**, 19 499–19 514.
- Weaver, A. T. and P. Courtier, 2001: Correlation modelling on the sphere using a generalized diffusion equation. *Q.J.R. Meteorol. Soc.*, **127 (575)**, 1815–1846.
- Weaver, A. T., C. Deltel, E. Machu, S. Ricci, and N. Daget, 2005: A multivariate balance operator for variational ocean data assimilation. *Q.J.R. Meteorol. Soc.*, **131 (613)**, 3605–3625.
- Weaver, A. T. and I. Mirouze, 2012: On the diffusion equation and its application to isotropic and anisotropic correlation modelling in variational assimilation. *Q.J.R. Meteorol. Soc.*, n/a–n/a.
- Weaver, A. T. and A. Piacentini, 2012: Representation of correlated observation error in the absolute dynamic topography: theory and numerical implementation in NEMOVAR. *VODA report*, **D2.2**, 1–9.
- Weston, P., 2011: Progress toward the implementation of correlated observation errors in 4D-Var. *Met Office Forecasting Research Technical report*, **560**, 1–33.
- Wiggins, S., 1992: *Chaotic transport in dynamical systems*. Springer-Verlag, New York.

Yaremchuk, M., D. Nechaev, and C. Pan, 2011: A hybrid background error covariance model for assimilating glider data into a coastal ocean model. *Mon. Wea. Rev.*, 110301130716023.

Zhang, F., M. Zhang, and J. A. Hansen, 2009: Coupling ensemble Kalman filter with four-dimensional variational data assimilation. *Adv. Atmos. Sci.*, **26 (1)**, 1–8.